

=> d his

FILE 'HCAPLUS' ENTERED AT 15:14:38 ON 29 DEC 1999

E LIPPS B/AU
L1 9 S E5,E9
E LIPPS F/AU
L2 6 S E3,E4,E6
L3 11 S L1-L2
L4 4 S LTNF
L5 29 S LETHAL? (L) TOXIN? (L) NEUTRALI? (L) FACTOR

Point of Contact:
Jan Delaval
Librarian-Physical Sciences
CM1 1E01 Tel: 308-4498

FILE 'REGISTRY' ENTERED AT 15:17:52 ON 29 DEC 1999

L6 10 S E1-E10
L7 4 S L6 AND SQL/FA
L8 3 S L7 NOT MAN/CI
E LKAMD/SQEP
L9 1 S E3
E LKAMDPTPPL/SQEP
L10 1 S E3
E LKAMDPTPPLWIKTE/SQEP
L11 1 S E3
L12 3 S L9-L11

FILE 'HCAPLUS' ENTERED AT 15:21:07 ON 29 DEC 1999

L13 2 S L12
L14 3 S L3 AND L4,L5
L15 2 S L3 AND L13
L16 3 S L13-L15
L17 3 S LETHAL? TOXIN? NEUTRAL? FACTOR
L18 0 S ANTILETHAL? TOXIN? NEUTRAL? FACTOR
L19 0 S ANTILTNF
L20 3 S L16,L17

FILE 'BIOSIS' ENTERED AT 15:25:26 ON 29 DEC 1999

E LIPPS B/AU
L21 9 S E3,E6,E7
E LIPPS F/AU
L22 0 S LTNF
L23 0 S L17
L24 0 S L12
L25 0 S L21 AND ?TOXIN?
L26 0 S L21 AND OPPOSUM

FILE 'WPIDS' ENTERED AT 15:27:05 ON 29 DEC 1999

L27 0 S LNTF OR ANTILNTF
L28 2 S LETHAL? TOXIN? NEUTRAL? FACTOR
E LIPPS B/AU
L29 5 S E5,E8
L30 5 S L28,L29

=> fil reg

FILE 'REGISTRY' ENTERED AT 15:32:32 ON 29 DEC 1999

USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.

PLEASE SEE "HELP USAGETERMS" FOR DETAILS.

COPYRIGHT (C) 1999 American Chemical Society (ACS)

STRUCTURE FILE UPDATES: 28 DEC 99 HIGHEST RN 251933-56-9

DICTIONARY FILE UPDATES: 28 DEC 99 HIGHEST RN 251933-56-9

TSCA INFORMATION NOW CURRENT THROUGH JANUARY 13, 1999

Please note that search-term pricing does apply when
conducting SmartSELECT searches.

POTENTIAL STEREO BOND SEARCH PROBLEM WITH STN EXPRESS WITH DISCOVER!

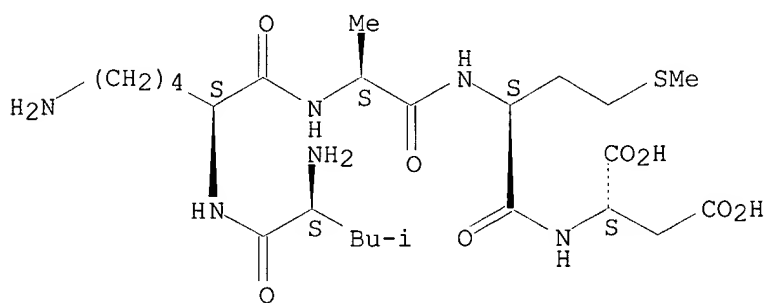
5.0 (Windows Only) SEE NEWS MESSAGE FOR DETAILS.

=> d sqide can tot l12

L12 ANSWER 1 OF 3 REGISTRY COPYRIGHT 1999 ACS
 RN 206755-64-8 REGISTRY
 CN L-Aspartic acid, L-leucyl-L-lysyl-L-alanyl-L-methionyl- (9CI) (CA INDEX NAME)
 FS PROTEIN SEQUENCE; STEREOSEARCH
 SQL 5

SEQ 1 LKAMD
 =====
 HITS AT: 1-5
 MF C24 H44 N6 O8 S
 SR CA
 LC STN Files: CA, CAPLUS, TOXLIT, USPATFULL

Absolute stereochemistry.



1 REFERENCES IN FILE CA (1967 TO DATE)
 1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

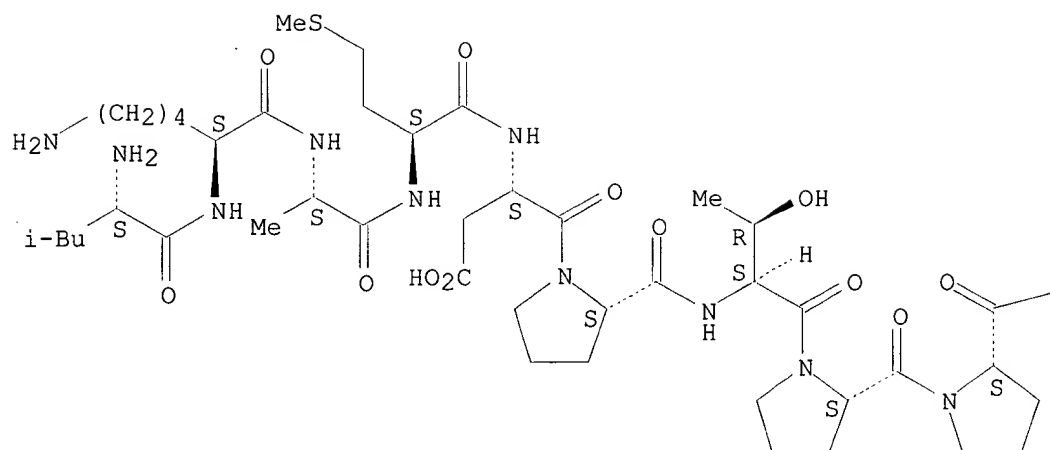
REFERENCE 1: 128:305070

L12 ANSWER 2 OF 3 REGISTRY COPYRIGHT 1999 ACS
 RN 206755-63-7 REGISTRY
 CN L-Leucine, L-leucyl-L-lysyl-L-alanyl-L-methionyl-L-.alpha.-aspartyl-L-prolyl-L-threonyl-L-prolyl-L-prolyl- (9CI) (CA INDEX NAME)
 FS PROTEIN SEQUENCE; STEREOSEARCH
 SQL 10

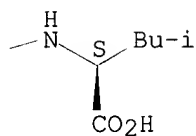
SEQ 1 LKAMDPTPPL
 =====
 HITS AT: 1-10
 MF C49 H83 N11 O14 S
 SR CA
 LC STN Files: CA, CAPLUS, TOXLIT, USPATFULL

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



1 REFERENCES IN FILE CA (1967 TO DATE)
 1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

REFERENCE 1: 128:305070

L12 ANSWER 3 OF 3 REGISTRY COPYRIGHT 1999 ACS
 RN 184782-42-1 REGISTRY
 CN L-Glutamic acid, L-leucyl-L-lysyl-L-alanyl-L-methionyl-L-.alpha.-aspartyl-
 L-prolyl-L-threonyl-L-prolyl-L-prolyl-L-leucyl-L-tryptophyl-L-isoleucyl-L-
 lysyl-L-threonyl- (9CI) (CA INDEX NAME)
 FS PROTEIN SEQUENCE; STEREOSEARCH
 SQL 15

SEQ 1 LKAMDPTPPL WIKTE

HITS AT: 1-15

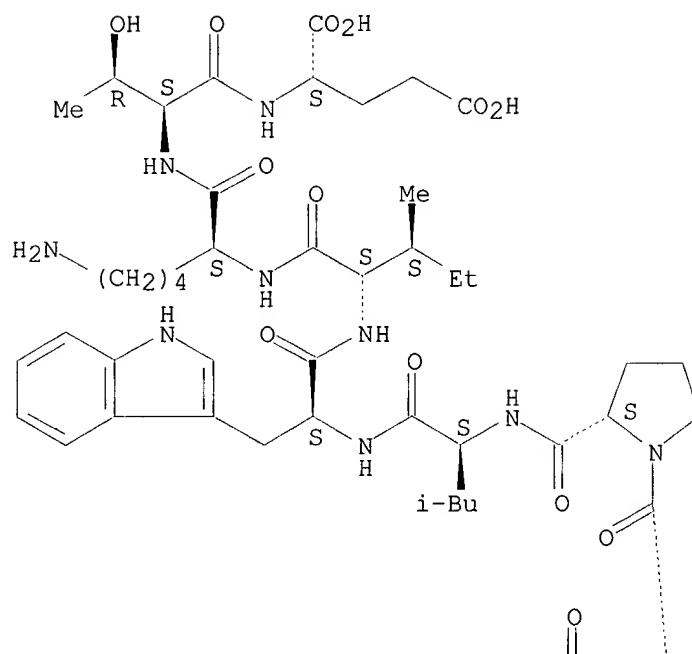
MF C81 H130 N18 O22 S

SR CA

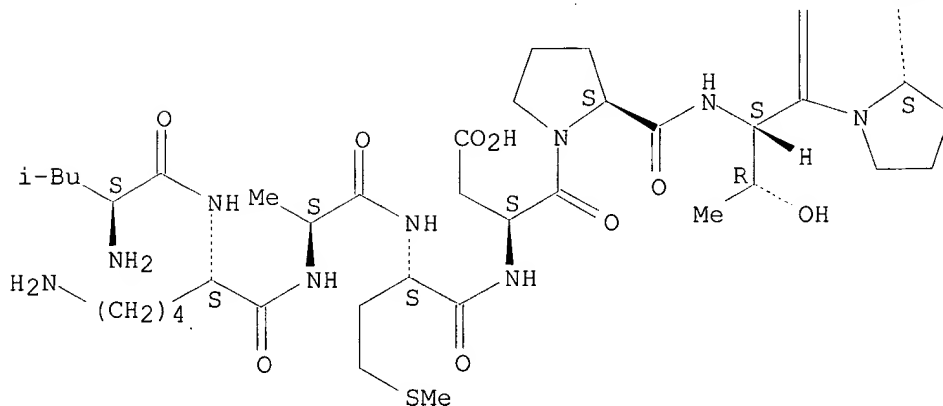
LC STN Files: CA, CAPLUS, TOXLIT, USPATFULL

Absolute stereochemistry.

PAGE 1-A



PAGE 2-A



2 REFERENCES IN FILE CA (1967 TO DATE)
 2 REFERENCES IN FILE CAPLUS (1967 TO DATE)

REFERENCE 1: 128:305070

REFERENCE 2: 126:27960

=> fil hcaplus

FILE 'HCAPLUS' ENTERED AT 15:32:48 ON 29 DEC 1999

USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.

PLEASE SEE "HELP USAGETERMS" FOR DETAILS.

COPYRIGHT (C) 1999 AMERICAN CHEMICAL SOCIETY (ACS)

Copyright of the articles to which records in this database refer is held by the publishers listed in the PUBLISHER (PB) field (available for records published or updated in Chemical Abstracts after December

26, 1996), unless otherwise indicated in the original publications.

FILE COVERS 1967 - 29 Dec 1999 VOL 132 ISS 1
FILE LAST UPDATED: 28 Dec 1999 (19991228/ED)

This file contains CAS Registry Numbers for easy and accurate substance identification.

This file supports REGISTRY for direct browsing and searching of all substance data from the REGISTRY file. Enter HELP FIRST for more information.

=> d all tot 120

L20 ANSWER 1 OF 3 HCAPLUS COPYRIGHT 1999 ACS
AN 1999:268585 HCAPLUS
DN 131:112584
TI Anti-lethal factor from opossum serum is a potent antidote for animal,
plant and bacterial toxins
AU **Lipps, B. V.**
CS Ophidia Products, Inc., Houston, TX, USA
SO J. Venomous Anim. Toxins (1999), 5(1), 56-66
CODEN: JVTOFG; ISSN: 0104-7930
PB Center for the Study of Venoms and Venomous Animals
DT Journal; (computer magnetic disk)
LA English
CC 4-5 (Toxicology)
AB Currently, the use of antivenoms is the only available treatment for
envenomation caused by venomous animals namely, snake, scorpion, spider,
tick and jelly fish. Antivenoms are generally produced in large animals,
mostly in horses. A large percentage of the population is allergic to
horse proteins. Several animals are known to be resistant to snakebites
and the antihemorrhagic and anti-lethal components have been
isolated from sera of opossum, mongoose, meerkat and hedgehog, as well as
from venomous and non-venomous snakes. Anti-lethal
factor named lethal toxin neutralizing
factor (LTNF) has been isolated in purity from opossum
(Didelphis virginiana) serum by high pressure liq. chromatog. (HPLC). The
mol. wt. of **LTNF** is 63 kDa, and it does not form pptn. with
venoms or **toxins** by immunodiffusion. Death due to i.p. (IP)
injection of a predetd. LD of venom from major families of snakes, for
instance Crotalidae, Elapidae, Viperidae and Hydrophiidae, is prevented in
mice by subsequent IP inoculation of **LTNF**. Furthermore,
LTNF neutralizes the lethality of scorpion and
bee venoms and **toxins** from various animals, plants and bacteria.
Thus, natural **LTNF** from opossum serum has potential as a
universal therapy for envenomation caused by animals, plants and bacteria.
ST opossum blood lethal factor antidote toxin; Didelphis blood lethal factor
antidote toxin
IT Antidotes
Didelphis virginiana
Poisoning (biological)
Serum (blood)
Snake venoms
Venoms
(antilethal factor from opossum serum is potent antidote for animal,
plant and bacterial toxins)
IT Ricins
Toxins
RL: ADV (Adverse effect, including toxicity); BIOL (Biological study)
(antilethal factor from opossum serum is potent antidote for animal,
plant and bacterial toxins)
IT Antitoxins
RL: THU (Therapeutic use); BIOL (Biological study); USES (Uses)
(antilethal factor from opossum serum is potent antidote for animal,

plant and bacterial toxins)
 IT 9007-40-3, Crotoxin 11029-72-4, Holothurin 12584-83-7, Cobratoxin
 52019-39-3, Taipoxin 107231-12-9, Botulin
 RL: ADV (Adverse effect, including toxicity); BIOL (Biological study)
 (antilethal factor from opossum serum is potent antidote for animal,
 plant and bacterial toxins)
 IT 9001-84-7, Phospholipase a2
 RL: BAC (Biological activity or effector, except adverse); BIOL
 (Biological study)
 (antilethal factor from opossum serum is potent antidote for animal,
 plant and bacterial toxins)

L20 ANSWER 2 OF 3 HCAPLUS COPYRIGHT 1999 ACS
 AN 1998:263232 HCAPLUS
 DN 128:305070
 TI **Lethal toxin neutralizing factors**
 IN **Lipps, Binie V.; Lipps, Frederick W.**
 PA USA
 SO U.S., 11 pp. Cont.-in-part of U.S. 5,576,297.
 CODEN: USXXAM
 DT Patent
 LA English
 IC ICM A61K038-10
 ICS A61K038-16; C07K007-08; C07K014-47
 NCL 514014000
 CC 4-5 (Toxicology)
 FAN.CNT 2

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	US 5744449	A	19980428	US 1996-657163	19960603
	US 5576297	A	19961119	US 1994-310340	19940922
PRAI	US 1993-58387		19930510		
	US 1994-310340		19940922		

AB Opossum whole serum exhibits a life saving property by neutralizing the lethality of venoms from all major families of poisonous snakes, and therefore an injection of Opossum serum can used as a novel treatment for many types of envenomation. Preferably, the injectable treatment for envenomation should be a compn. obtained from the fraction of Opossum whole serum which contains the **LTNF**, i. e. the so called "**LTNF-n**", in purity. A method is given for the manuf. of a **LTNF** from the serum of an opossum (*Didelphis virginiana*) serum, by fractionating the opossum serum and isolating this select fraction from the plurality of fractions having an N terminal amino acid sequence given by SEQ ID No: 1. A short peptide was synthesized having SEQ ID No: 1. The synthetic peptide having the sequence SEQ ID No: 1 shows lethal toxin neutralizing activity similar to the natural **LTNF** from opossum or mongoose sera. The synthetic **LTNF** also has life saving utility.

ST opossum serum **lethal toxin neutralizing factor**

IT Poisoning (biological)
 RL: ADV (Adverse effect, including toxicity); BIOL (Biological study)
 (botulism; opossum serum-derived **lethal toxin neutralizing factor** for universal treatment of snakebite or other envenomation and intoxication)

IT Peptides, biological studies
 RL: PRP (Properties); PUR (Purification or recovery); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
 (**lethal toxin neutralizing factors** (250-2,500 mol. wt.); opossum serum-derived **lethal toxin neutralizing factor** for universal treatment of snakebite or other envenomation and intoxication)

IT Allergies
 Biological warfare agents
 Chemical warfare agents

Crotalus adamanteus
 Crotalus atrox
 Didelphidae
 Didelphis virginiana
 Elapidae
 Hydrophidae
 Inflammation
 Naja naja atra
 Naja naja kaouthia
 Oxyuranus scutellatus
 Protein sequences
 Sepsis
 Serum (blood)
 Venoms
 Vipera russelli
 Viperidae
 (opossum serum-derived **lethal toxin**
 neutralizing factor for universal treatment of
 snakebite or other envenomation and intoxication)

IT Ricins
 RL: ADV (Adverse effect, including toxicity); BIOL (Biological study)
 (opossum serum-derived **lethal toxin**
 neutralizing factor for universal treatment of
 snakebite or other envenomation and intoxication)

IT **Toxins**
 RL: ADV (Adverse effect, including toxicity); BIOL (Biological study)
 (plant or bacterial; opossum serum-derived **lethal**
 toxin neutralizing factor for universal
 treatment of snakebite or other envenomation and intoxication)

IT Snake
 (poisonous; opossum serum-derived **lethal toxin**
 neutralizing factor for universal treatment of
 snakebite or other envenomation and intoxication)

IT Bee
 Scorpion
 (sting; opossum serum-derived **lethal toxin**
 neutralizing factor for universal treatment of
 snakebite or other envenomation and intoxication)

IT Plant (Embryophyta)
 (**toxin**; opossum serum-derived **lethal toxin**
 neutralizing factor for universal treatment of
 snakebite or other envenomation and intoxication)

IT Bacteria (Eubacteria)
 RL: ADV (Adverse effect, including toxicity); BIOL (Biological study)
 (**toxin**; opossum serum-derived **lethal toxin**
 neutralizing factor for universal treatment of
 snakebite or other envenomation and intoxication)

IT 9001-84-7, Phospholipase A2 9007-40-3, Crotoxin 12584-83-7, Cobratoxin
 52019-39-3, Taipoxin
 RL: ADV (Adverse effect, including toxicity); BIOL (Biological study)
 (opossum serum-derived **lethal toxin**
 neutralizing factor for universal treatment of
 snakebite or other envenomation and intoxication)

IT 51-45-6, Histamine, biological studies
 RL: ADV (Adverse effect, including toxicity); BSU (Biological study,
 unclassified); BIOL (Biological study)
 (opossum serum-derived **lethal toxin**
 neutralizing factor for universal treatment of
 snakebite or other envenomation and intoxication)

IT **184782-42-1 206755-63-7 206755-64-8**
 RL: PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES
 (Uses)
 (opossum serum-derived **lethal toxin**
 neutralizing factor for universal treatment of
 snakebite or other envenomation and intoxication)

L20 ANSWER 3 OF 3 HCAPLUS COPYRIGHT 1999 ACS

AN 1996:713649 HCAPLUS

DN 126:27960

TI Embodiments of natural and synthetic **lethal toxin neutralizing factors** and their utility as treatment for envenomation

IN **Lipps, Binie V.; Lipps, Frederick W.**

PA Lipps; Binie V., USA; Lipps; Frederick W.

SO U.S., 9 pp. Cont.-in-part of U.S. Ser. No. 53,387, abandoned.

CODEN: USXXAM

DT Patent

LA English

IC ICM A61K038-10

ICS A61K038-16; C07K007-08; C07K014-47

NCL 514014000

CC 4-5 (Toxicology)

FAN.CNT 2

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	US 5576297	A	19961119	US 1994-310340	19940922
	US 5744449	A	19980428	US 1996-657163	19960603
PRAI	US 1993-58387		19930510		
	US 1994-310340		19940922		

AB Opossum whole serum exhibits a life saving property by **neutralizing** the **lethality** of venoms from all major families of poisonous snakes, and therefore an injection of Opossum serum can be used as a novel treatment for many types of envenomation. Preferably, the injectable treatment for envenomation should be a compn. obtained from the fraction of Opossum whole serum which contains the **lethal toxin neutralizing factor**, i.e. the so called natural **LTNF**, in purity. A method is given for the manuf. of a **lethal toxin neutralizing factor** from the serum of an opossum (*Didelphis virginiana*), by fractionating the opossum serum and isolating this select fraction from the plurality of fractions having an N terminal amino acid sequence given by SEQ ID No: 1. A short peptide was synthesized having SEQ ID No: 1. The synthetic peptide having sequence SEQ ID No: 1 shows **lethal toxin neutralizing** activity similar to the natural **LTNF** from opossum or mongoose sera. The synthetic **LTNF** also has life saving utility.

ST opossum **lethal toxin neutralizing factor** envenomation; *Didelphis* **lethal toxin neutralizing factor** envenomation

IT Coagulation **factors** (blood)
 RL: PUR (Purification or recovery); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
 (**lethal toxin neutralizing factor** ; natural and synthetic **lethal toxin neutralizing factors** and utility as treatment for envenomation)

IT *Didelphis virginiana*
 Serum (blood)
 (**lethal toxin neutralizing factors** from opossum serum and utility as treatment for envenomation)

IT Antivenoms
 Snake
 (natural and synthetic **lethal toxin neutralizing factors** and utility as treatment for envenomation)

IT **184782-42-1P**
 RL: PUR (Purification or recovery); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
 (natural and synthetic **lethal toxin neutralizing factors** and utility as treatment for

envenomation)

=> fil wpids

FILE 'WPIDS' ENTERED AT 15:33:02 ON 29 DEC 1999
COPYRIGHT (C) 1999 DERWENT INFORMATION LTD

FILE LAST UPDATED: 21 DEC 1999 <19991221/UP>

>>>UPDATE WEEKS:

MOST RECENT DERWENT WEEK 199954 <199954/DW>
DERWENT WEEK FOR CHEMICAL CODING: 199954
DERWENT WEEK FOR POLYMER INDEXING: 199954
DERWENT WORLD PATENTS INDEX SUBSCRIBER FILE, COVERS 1963 TO DATE

>>> D COST AND SET NOTICE DO NOT REFLECT SUBSCRIBER DISCOUNTS -
SEE HELP COST <<<

>>> IMPORTANT DERWENT ANNOUNCEMENT ABOUT CHANGES TO CPI
SUBSCRIBER INDEXING - SEE NEWS <<<

>>> FOR UP-TO-DATE INFORMATION ABOUT ALL 'NEW CONTENT' CHANGES TO
WPIDS, INCLUDING THE DERWENT CHEMISTRY RESOURCE (DCR),
PLEASE VISIT <http://www.derwent.com/newcontent.html> <<<

=> d all abeq tot 130

L30 ANSWER 1 OF 5 WPIDS COPYRIGHT 1999 DERWENT INFORMATION LTD

AN 1998-271108 [24] WPIDS

CR 1997-011287 [01]

DNC C1998-084507

TI **Lethal Toxin Neutralising Factor**

peptide from opossum - can neutralise venom(s) from all major families of
poisonous snakes.

DC B04

IN **LIPPS, B V; LIPPS, F W**

PA (LIPP-I) LIPPS B V; (LIPP-I) LIPPS F W

CYC 1

PI US 5744449 A 19980428 (199824)* 11p A61K038-10

ADT US 5744449 A CIP of US 1993-58387 19930510, CIP of US 1994-310340
19940922, US 1996-657163 19960603

FDT US 5744449 A CIP of US 5576297

PRAI US 1996-657163 19960603; US 1993-58387 19930510; US 1994-310340
19940922

IC ICM A61K038-10

ICS A61K038-16; C07K007-08; C07K014-47

AB US 5744449 A UPAB: 19980617

A new peptide (I) has a molecular weight of 250 to 2500 g/mole and
comprises at least a 3 amino acid portion of:

Leu-Lys-Ala-Met-Asp-Pro-Thr-Pro-Pro-Leu-Trp-Ile-Lys-Thr-Glu (A).

Also claimed are:

(1) a peptide as (I), but comprising at least a 5 amino acid portion
of (A);

(2) a method for treating a victim of envenomation from a poisonous
snake comprising administering (I) to a victim by a route and in an amount
which is effective to neutralise at least one effect of the envenomation;
and

(3) a method for treating a victim of envenomation from a poisonous
snake family from the family of Elapidae, Viperidae or sea snake by
administering a protein of about 68 kDa with an initial 15 amino acid
sequence as (A) by a route and in an amount which is effective to
neutralise at least one effect of the envenomation.

USE - The peptide of (I) and the methods is a **Lethal
Toxin Neutralising Factor** (LTNF) moiety from a
68 kDa anti-haemorrhagic protein derived from an opossum. It is useful for
the treatment of snake bites, sepsis, allergies caused by the environment

and treatment of bee or scorpion stings or toxicities caused by plant or bacterial toxins. (I) can also be used in histamine reaction treatment (claimed).

ADVANTAGE - The peptides can be used in envenomation treatment for a variety of snakes without prior identification of the guilty snake. Being short it can be synthetically prepared rather than the current production in horses, where some people can show hypersensitivity to horse proteins.

Dwg.0/3

FS CPI
FA AB; DCN
MC CPI: B04-C01C; B04-N04A; B14-A01; B14-G02A; B14-M01

L30 ANSWER 2 OF 5 WPIDS COPYRIGHT 1999 DERWENT INFORMATION LTD
AN 1997-372086 [34] WPIDS
DNC C1997-119850
TI Topical treatment of herpes simplex virus infections - with combination of snake venom proteins.
DC B04
IN **LIPPS, B V; LIPPS, F W**
PA (LIPP-I) LIPPS B V; (LIPP-I) LIPPS F W
CYC 1
PI US 5648339 A 19970715 (199734)* 8p A61K038-16
ADT US 5648339 A US 1994-280157 19940725
PRAI US 1994-280157 19940725
IC ICM A61K038-16
ICS A61K035-58; A61K038-17
AB US 5648339 A UPAB: 19970820
Treatment of herpes simplex virus (HSV-1 or HSV-2) infections comprises topically applying to an affected area a composition comprising two phospholipase A2 proteins obtained from the snake venom of Naja naja kaouthia. The proteins have molecular weights of 14 and 15 kD and have the N-terminal amino acid sequence Asn-Leu-Tyr-Gln-Phe-Lys-Asn-Met-Ile-Gln-Cys-Thr-Val-Pro-Asn.

In an example, Chang's liver cell cultures were infected with HSV-1 or HSV-2 at 10⁻² to 10⁻⁶ dilutions. After an absorption period, excess inoculum was removed and 1 ml of medium was added to each well. A 1:1 mixture of the two proteins was added at a concentration of 5 µg/ml. In all cases, no cytopathic effect was observed after 4 days.

ADVANTAGE - The composition inhibits the cytopathic effect of HSV-1 and HSV-2 at concentrations as low as 5 µg/ml without being cytotoxic.

Dwg.0/3

FS CPI
FA AB; DCN
MC CPI: B04-N02; B14-A02A3

L30 ANSWER 3 OF 5 WPIDS COPYRIGHT 1999 DERWENT INFORMATION LTD
AN 1997-011287 [01] WPIDS
CR 1998-271108 [24]
DNC C1997-003058
TI Treatment of victims of bee or scorpion stings or plant or bacterial toxins - by admin. of **lethal toxin-neutralising factor** or its N-terminal peptide.
DC B04 D16
IN **LIPPS, B V; LIPPS, F W**
PA (LIPP-I) LIPPS B V; (LIPP-I) LIPPS F W
CYC 1
PI US 5576297 A 19961119 (199701)* 9p A61K038-10
ADT US 5576297 A CIP of US 1993-58387 19930510, US 1994-310340 19940922
PRAI US 1994-310340 19940922; US 1993-58387 19930510
IC ICM A61K038-10
ICS A61K038-16; C07K007-08; C07K014-47
AB US 5576297 A UPAB: 19980617
Methods for treating victims of bee stings, scorpion stings, plant toxins or bacterial toxins comprise intravenous injection of (a) a **lethal toxin-neutralising factor** (LTNF) obtained from an animal having resistance to envenomation, provided that the animal

is a Didelphis opossum in the case of bee and scorpion stings, or (b) a peptide of formula (I): Leu-Lys-Ala-Met-Asp-Pro-Thr-Pro-Pro-Leu-Trp-Ile-Lys-Thr-Glu (I). Also claimed is a method for treating a victim of a bee sting, comprising topical admin. of a soln. of (I).

ADVANTAGE - Use of natural or synthetic LTNF as a treatment for snake bites overcomes the problem of hypersensitivity occurring with horse-derived antivenom.

Dwg.0/3

FS CPI

FA AB; DCN

MC CPI: B04-C01C; B04-H01; B14-G02A; B14-M01; D05-H13

L30 ANSWER 4 OF 5 WPIDS COPYRIGHT 1999 DERWENT INFORMATION LTD

AN 1996-476328 [47] WPIDS

DNC C1996-148746

TI Treatment of cancers - using Atroporin or Kaotree or a combination, which are purified from the venom of snakes..

DC B04

IN **LIPPS, B V; LIPPS, F W**

PA (LIPP-I) LIPPS B V; (LIPP-I) LIPPS F W

CYC 1

PI US 5565431 A 19961015 (199647)* 20p A61K038-17

ADT US 5565431 A CIP of US 1994-262678 19940620, US 1994-310341 19940922

PRAI US 1994-310341 19940922; US 1994-262678 19940620

IC ICM A61K038-17

ICS C07K014-46

AB US 5565431 A UPAB: 19961124

A claimed method for treating a patient having a cancer tumour comprising cancer cells comprises contacting the cancer cells with molecules of Atroporin and/or Kaotree at a dosage level sufficient to slow or regress the cancer tumour without significant destruction of normal cells which may also be contacted by Atroporin and/or Kaotree. The Atroporin and Kaotree each consist of a single purified fraction obtd. from snake venom, each fraction contg. a single peptide having either its first fifteen N-terminal amino acids as in (I) and a mol. wt. of about 35000 Daltons (for Atroporin) or having its first fifteen N-terminal amino acids as in (II) and a mol. wt. of about 6000 Daltons (for Kaotree).

Xaa-Xaa-Gly-Gly-Asp-Glu-Cys-Asn-Ile-Asn-Glu-His-Arg-Ser-Leu (I)

Met-Glu-Cys-Tyr-Arg-Met-Ser-Asn-Ile-Val-Thr-Cys-Gln-Pro-Trp (II)

USE - Atroporin and/or Kaotree show cytolytic activity against a wide range of tumour cells and so are useful as anticancer agents against e.g. cancers of the breast, colon, liver, lung, pharynx and ovary.

ADVANTAGE - The peptides do not cause the typical side effects caused by usual chemotherapy.

Dwg.0/7

FS CPI

FA AB; DCN

MC CPI: B04-N02B; B14-H01B

L30 ANSWER 5 OF 5 WPIDS COPYRIGHT 1999 DERWENT INFORMATION LTD

AN 1995-393068 [50] WPIDS

DNC C1995-169353

TI Use of beta-taipoxin as cell growth factor and mitogen - partic. to promote cell growth in serum free media and to improve healing of wounds.

DC B04 D16

IN **LIPPS, B V**

PA (LIPP-I) LIPPS B V

CYC 61

PI WO 9529987 A1 19951109 (199550)* EN 20p C12N005-06

RW: AT BE CH DE DK ES FR GB GR IE IT KE LU MC MW NL OA PT SD SE SZ UG

W: AM AU BB BG BR BY CA CN CZ FI GE HU JP KG KP KR KZ LK LR LT LV MD

MG MN MX NO NZ PL RO RU SI SK TJ TT UA US UZ VN

AU 9526360 A 19951129 (199609) C12N005-06

EP 759069 A1 19970226 (199714) EN C12N005-06

R: BE CH DE FR GB IT LI SE

AU 691621 B 19980521 (199832) C12N005-06

ADT WO 9529987 A1 WO 1995-US5657 19950503; AU 9526360 A AU 1995-26360
19950503; EP 759069 A1 EP 1995-921232 19950503, WO 1995-US5657 19950503;
AU 691621 B AU 1995-26360 19950503

FDT AU 9526360 A Based on WO 9529987; EP 759069 A1 Based on WO 9529987; AU
691621 B Previous Publ. AU 9526360, Based on WO 9529987

PRAI US 1994-237129 19940503

REP 06Jnl.Ref; US 4443546

IC ICM C12N005-06

ICS A61K038-18; C12N005-08

AB WO 9529987 A UPAB: 19951215

A culture medium contg. a peptide (I) having the N-terminal sequence (Ia)
is new Asn-Leu-Val-Glu-Phe-Gly-Lys-Met-Ile-Glu-Cys-Ala-Ile-Arg-Asn
(Ia).

USE - (I) is partic. beta-taipoxin isolated from snake venom. (I) is
a cell growth factor and potent mitogen. It provides good cell growth in
absence of serum (e.g. for growth of skin monolayers) and provides rapid,
scar-free healing of cuts, burns and abrasions (when applied topically).

ADVANTAGE - Cells can now be grown effectively in serum-free medium
so purification of cellular products is facilitated. At 0.1 mug/ml (Ia)
has an effect equiv. to 10% serum for a wide range of cells. (Ia) is water
soluble, stable at 4deg.C or room temp. and not toxic to cells at concns.
of use.

Dwg.0/3

FS CPI

FA AB; DCN

MC CPI: B04-C01C; B04-F02; B14-N17B; D05-H08; D05-H17A2

=> d his 131-

FILE 'WPIDS' ENTERED AT 15:33:02 ON 29 DEC 1999

L31 2 S LEU LYS ALA MET ASP PRO THR PRO PRO LEU TRP ILE LYS THR GLU

L32 0 S L31 NOT L30

Wed Dec 29 14:17:34 1999

baskar-09-300612.pcp

Page 1

: Entered [jdelaval 29-Dec-99 14:04]
09-300612
lkamdpctppiwktel

MIPSOL
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mparch_dp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 29 15:52:08 1999; MasPar time 8.77 Seconds

Tabular output not generated. 88.426 Million cell updates/sec

Title: >09-300612
Description: (1-10) from baskar-09-300612.pep
Perfect Score: 73
Sequence: 1 lkamdptppl 10

Scoring table: PAM 150
Gap 15

Searched: 547422 seqs, 77552421 residues

Post-processing: Minimum Match 100%
Listing first 1000 summaries
Maximum DB seq length 10

Database: a-pending
1:P9 2:U60 3:U7 4:U80 5:U81 6:U82 7:U83 8:U84A 9:U84B
10:U85 11:U86 12:U87 13:U88 14:U89 15:U90 16:U91 17:U92
18:U93 19:NEWP 20:NEWU6 21:NEWU7 22:NEWU8 23:NEWU9

Statistics: Mean 17.858; Variance 49.093; scale 0.364

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	No.	Score	Match	length	ID	Description	Pred.	No.

No matches found.

Search completed: Wed Dec 29 15:53:52 1999

Job time : 104 secs.

WIDEORLH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 29 15:45:18 1999; MasPar time 3.69 Seconds

Tabular output not generated. 108.566 Million cell updates/sec

Title: >09-300612

Description: (1-10) from baskar-09-300612.pep

Perfect Score: 73

Sequence: 1 lkamdptppl 10

Scoring table: PAM 150

Gap 15

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 100%

Listing first 1000 summaries

Maximum DB seq length 10

Database:

pir60
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 23.017; Variance 29.210; scale 0.788

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	
No.	Score	Match Length DB ID Description Pred. No.

No matches found.

Search completed: Wed Dec 29 15:46:16 1999
Job time : 58 secs.

M Q E N L (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 29 15:48:23 1999; MasPar time 2.76 Seconds
102.460 Million cell updates/sec

Tabular output not generated.

Title: >09-300612
Description: (1-10) from baskar-09-300612.pep
Percent Score: 73
Sequence: 1 lkamdppl 10

Scoring table: PAM 150
Gap 15

Searched: 77977 segs, 28268293 residues

Post-processing: Minimum Match 100%
Listing first 1000 summaries
Maximum DB seq length 10

Database: swiss-prot37
1:swissprot

Statistics: Mean 23.724; Variance 26.440; scale 0.897

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query					
No.	Score	Match	Length	ID	Description	Pred. No.

No matches found.

Search completed: Wed Dec 29 15:49:14 1999
Job time : 51 secs.

WISNTH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msearch_pp Protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 29 15:49:32 1999; MasPar time 5.28 Seconds
103.294 Million cell updates/sec

Tabular output not generated.

Title: >09-300612
Description: (1-10) from baskar-09-300612.pep
Perfect Score: 73
Sequence: 1 lkamdpppl 10

Scoring table: PAM 150
Gap 15

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 100%
Listing first 1000 summaries
Maximum DB seq length 10

Database: sptrembl9
1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
9:sp_phase 10:sp_plant 11:sp_rodent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

Statistics: Mean 22.746; Variance 27.733; scale 0.820

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description	Pred. No.

No matches found.

Search completed: Wed Dec 29 15:50:47 1999
Job time : 75 secs.

 WIDEORLH (TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 29 15:45:23 1999; Maspar time 3.21 Seconds

Tabular output not generated. 33.166 Million cell updates/sec

Title: >09-300612

Description: (1-5) from basket-09-300612.pep

Perfect Score: 36

Sequence: 1 lkamd 5

Scoring table: PAM 150

Gap 15

Searched: 170751 seqs, 21266608 residues

Post-processing: Minimum Match 100%

Listing first 1000 summaries

Maximum DB seq length 5

Database:

a-geneseg35
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29 30:part30 31:part31 32:part32 33:part33
 34:part34 35:part35 36:part36 37:part37 38:part38
 39:part39

Statistics: Mean 13.108; Variance 33.346; scale 0.393

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description	Pred. No.
1	36	100.0	5	30	W53842 N-terminus of opossum	3.31e+02

Note: Post-processor removed 999 summaries from list due to search parameters chosen.

ALIGNMENTS

RESULT 1
 ID W53842 standard; peptide; 5 AA.
 AC W53842;
 DT 08-JUL-1998 (first entry)
 DE N-terminus of opossum LTNF.
 KW LTNF, lethal toxin neutralising factor; Opossum; envenomation; therapy; anti-haemorrhagic protein; Elapidae; Viperidae; sea snake; snake bite; sepsis; allergy; bee sting; scorpion sting; plant toxin; bacterial toxin;

KW histamine reaction treatment.
 OS Didelphis virginiana.
 PN US5744449-A.
 PD 28-APR-1998.
 PF 03-JUN-1996; 657163.
 PR 10-MAY-1993; US-657163.
 PR 22-SEP-1994; US-310340.
 PA (LIPP/) LIPPS B V.
 PA (LIPP/) LIPPS F W.
 PI LIPPS BV, LIPPS FW;
 DR WPI: 98-271108/24.
 PT Lethal Toxin Neutralising Factor peptide from opossum - can
 PT neutralise venom(s) from all major families of poisonous snakes
 PS Claim 5, Column 13; LIPP, English.
 CC This sequence represents the peptide of the invention. It is a lethal
 CC Toxin Neutralising Factor (LTNF) moiety from a 68 kDa anti-haemorrhagic
 CC protein derived from an opossum. The peptide can be used in a method for
 CC treating a victim of envenomation from a poisonous snake, preferably a
 CC poisonous snake from the family of Elapidae, Viperidae or sea snake.
 CC It is useful for the treatment of snake bites, sepsis, allergies caused
 CC by the environment and treatment of bee or scorpion stings or toxicities
 CC caused by plant or bacterial toxins. The peptide can also be used in
 CC histamine reaction treatment. The peptide can be used in envenomation
 CC treatment for a variety of snakes without prior identification of the
 CC snake. Being short it can be synthetically prepared rather than the
 CC current production in horses, where some people can show hypersensitivity
 CC to horse proteins.
 SQ Sequence 5 AA;

Query Match 100.0%; Score 36; DB 30; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.31e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 lkamd 5
 QY 1 lkamd 5

Search completed: Wed Dec 29 15:46:41 1999
 Job time : 78 secs.

MIMOSA (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 29 15:50:54 1999; MasPar time 1.54 Seconds
38.739 Million cell updates/sec

Tabular output not generated.

Title: >09-300612

Description: (1-5) from baskar-09-300612.pep

Perfect Score: 36

Sequence: 1 lkamd 5

Scoring table: PAM 150
Gap 15

Searched: 122461 seqs, 11912985 residues

Post-processing: Minimum Match 100%

Listing first 1000 summaries

Maximum DB seq length 5

Database:

1:5A_COMB 2:5B_COMB 3:PCT9_COMB 4:backfile1

Statistics: Mean 12.206; Variance 31.995; scale 0.382

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	36	100.0	5	1	US-08-657- Sequence 3, Applcattio	1.73e+02

Note: Post-processor removed 822 summaries from list due to search parameters chosen.

RESULT 1 NOT AVAILABLE.

ALIGNMENTS

RESULT 1
ID US-08-657-163A-3 STANDARD; PRT; 5 AA.
AC xxxxxx
XX
XX
DT
XX
XX Sequence 3, Application US/08657163A
CC Sequence 3, Application US/08657163A
CC Patent No. 5744449
CC GENERAL INFORMATION:
CC APPLICANT: BINIE V. LIPPS AND FREDERICK W. LIPPS

CC TITLE OF INVENTION: EMBODIMENTS OF NATURAL AND
CC TITLE OF INVENTION: SYNTHETIC LIPNS AND THEIR
CC TITLE OF INVENTION: UTILITY AS TREATMENT FOR ENVENOMATION
CC NUMBER OF SEQUENCES: 3
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: BINIE V. LIPPS
CC STREET: 4509 MIMOSA DR.
CC CITY: BELLAIRE
CC STATE: TEXAS
CC COUNTRY: USA
CC ZIP: 77401
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5" FLOPPY DISK, 1.44 MB
CC COMPUTER: IBM COMPATIBLE
CC OPERATING SYSTEM: MS-DOS 5.0/WINDOWS 3.1
CC SOFTWARE: MS WORD 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/657,163A
CC FILING DATE:
CC CLASSIFICATION: 514
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/310,340
CC FILING DATE: 22 SEPTEMBER 1994
CC CLASSIFICATION: 514
CC APPLICATION NUMBER: 08/058,387
CC FILING DATE: 10 MAY 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: JOHN R. CASPERSON
CC REGISTRATION NUMBER: 28,198
CC REFERENCE/DOCKET NUMBER: FWL-PAT-US-011
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 713-482-2961
CC TELEFAX: 713-663-7290
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 5
CC TYPE: AMINO ACID
CC STRANDEDNESS: SINGLE
CC TOPOLOGY: LINEAR
CC MOLECULE TYPE: PEPTIDE IN SEQ ID NO: 3
CC HYPOTHETICAL: YES
CC ANTI-SENSE: NO
CC FRAGMENT TYPE: N
CC ORIGINAL SOURCE: SYNTHETIC
CC SEQUENCE 5 AA: 577 MW; 110 CN;
SQ
Query Match 100.0%; Score 36; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.73e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 LKAMD 5
Qy 1 lkamd 5

Search completed: Wed Dec 29 15:51:37 1999
Job time : 43 secs.

MIPS LIMITED (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPerch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 29 15:51:56 1999; MasPar time 7.13 Seconds

Tabular output not generated. 54.374 Million cell updates/sec

Title: >09-300612
Description: (1-5) from baskar-09-300612.pep
Perfect Score: 36
Sequence: 1_kamd 5

Scoring table: PAM 150
Gap 15

Searched: 547422 segs, 77552421 residues

Post-processing: Minimum Match 100%
Listing first 1000 summaries
Maximum DB seq length 5

Database: a-pending
1:P9 2:U60 3:U7 4:U80 5:U81 6:U82 7:U83 8:U84A 9:U84B
10:U85 11:U86 12:U87 13:U88 14:U89 15:U90 16:U91 17:U92
18:U93 19:NEWP 20:NEWU6 21:NEWU7 22:NEWU8 23:NEWU9

Statistics: Mean 14.425; Variance 31.327; scale 0.460

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query				
No.	Score	Match length	ID	Description	Pred. No.

No matches found.

Search completed: Wed Dec 29 15:54:54 1999
Job time : 178 secs.

M P E R C H
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mpsrch_dp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 29 15:46:58 1999; MasPar time 2.93 Seconds

Tabular output not generated. 68.364 Million cell updates/sec

Title: >09-300612

Description: (1-5) from baskar-09-300612.pep

Perfect Score: 36

Sequence: 1 lkamd 5

Scoring table: PAM 150
Gap 15

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 100%

Listing first 1000 summaries

Maximum DB seq length 5

Database:

pir60
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 18.837; Variance 18.467; scale 1.020

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description	Pred. No.

No matches found.

Search completed: Wed Dec 29 15:47:54 1999
Job time : 56 secs.

M I S R I H (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch.p protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 29 15:48:11 1999; MasPar time 2.03 Seconds

Tabular output not generated. 69.541 Million cell updates/sec

Title: >09-300612

Description: (1-5) from baskar-09-300612.pep

Perfect Score: 36

Sequence: 1 Ikamd 5

Scoring table: PAM 150
Gap 15

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 100%

Listing first 1000 summaries

Maximum DB seq length 5

Database: swiss-prot37
1:swissprot

Statistics: Mean 19.377; Variance 15.898; scale 1.219

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query					
No.	Score	Match	length	ID	Description	Pred. No.

No matches found.

Search completed: Wed Dec 29 15:48:59 1999
Job time : 48 secs.

M P E S R E L I
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPrch_PP protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 29 15:49:17 1999; MasPar time 4.11 Seconds

Tabular output not generated. 66.354 Million cell updates/sec

Title: >09-300612
Description: (1-5) from baskar-09-300612.pep
Perfect Score: 36
Sequence: 1 lkamd 5

Scoring table: PAM 150
Gap 15

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 100%
Listing first 1000 summaries
Maximum DB seq length 5

Database: spiremb19
1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
9:sp_phage 10:sp_plant 11:sp_rodent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

Statistics: Mean 18.536; Variance 16.495; scale 1.124

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query					
No.	Score	Match	Length	ID	Description	Pred. No.

No matches found.

Search completed: Wed Dec 29 15:50:37 1999
Job time : 80 secs.

 MIPSE

 (TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mpsrch_dp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 29 14:06:36 1999; MasPar time 2.57 Seconds

Tabular output not generated. 164,988 Million cell updates/sec

Title: >09-300612

Description: (1-15) from bascar-09-300612.pep

Sequence: 116
 1 lkamdpplwikte 15

Scoring table: PAM 150
 Gap 15

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: swiss-prot37
 1:swissprot

Statistics: Mean 27.347; Variance 37.123; scale 0.737

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	72	62.1	1213	1	T2D2_DROME TRANSCRIPTION INITIAT	6.17e-03
2	64	55.2	662	1	FAST1_SCHAM FASCICLIN I PRECURSOR	2.51e-01
3	62	53.4	271	1	THRR_AZOVI THIOSULFATE SULFOTRAN	6.08e-01
4	59	50.9	102	1	YCA_TYUCA HYPOHETICAL 11.4 KD P	2.22e+00
5	59	50.9	331	1	DIV_ECOLI DIV PROTEIN.	2.22e+00
6	59	50.9	564	1	NCAP_LYCVU NUCLEOCAPSID PROTEIN (2.22e+00
7	59	50.9	1227	1	YAVI_RHISN IAF4-4 PROTEIN (LYMPHOI	3.38e+00
8	58	50.0	351	1	YAVI_RHISN HYPOHETICAL 39.2 KD P	3.38e+00
9	58	50.0	1305	1	RRPL_AHSVS RNR-DIRECTED RNA POLYM	3.38e+00
10	58	50.0	1741	1	COA_HUMAN COMPLEMENT C4 PRECURSO	5.12e+00
11	57	49.1	209	1	YSCK_YERPS YOP PROTEINS TRANSLA	5.12e+00
12	57	49.1	305	1	YSCK_YEREN YOP PROTEINS TRANSLA	5.12e+00
13	57	49.1	355	1	REF_HAEIN PUTATIVE UNDECAPRENTY-	5.12e+00
14	56	48.3	192	1	YAPG_RHISN HYPOHETICAL 21.1 KD P	7.73e+00
15	56	48.3	328	1	SRG1_CAEEL SRC-1 PROTEIN	7.73e+00
16	56	48.3	517	1	CPN1_LANCA CYTOCHROME P450 XIB PR	7.73e+00
17	55	47.4	357	1	MITO_HUMAN L-MYC-2 PROTEIN.	1.16e+01
18	55	47.4	425	1	YOE_ECOLI HYPOHETICAL 46.8 KD P	1.16e+01
19	55	47.4	453	1	AATM_ARATH ASPARTATE AMINOTRANSF	1.16e+01
20	55	47.4	522	1	CYSJ_THIRO SULFITE REDUCTASE [NAN	1.16e+01
21	55	47.4	550	1	CBS_HUMAN CYSTATHIONINE BETA-SYN	1.16e+01
22	55	47.4	594	1	UVRC_HELPY EXCINUCLEASE ABC SUBUN	1.16e+01
23	55	47.4	679	1	TKT1_YEAST TRANSEKTOLEASE 1 (EC 2.	1.16e+01

24	55	47.4	865	1	ENV_SIVAT ENVELOPE POLYPROTEIN G	1.16e+01
25	55	47.4	932	1	HMDH_STRPU 3-HYDROXY-3-METHYLGLUT	1.16e+01
26	55	46.6	41	1	AIBG_HORSE ALPHA-1B-GLYCOPROTEIN	1.73e+01
27	54	46.6	171	1	TRAV_ECOLI TRAV PROTEIN PRECURSOR	1.73e+01
28	54	46.6	307	1	CD5R_MOUSE CYCLIN-DEPENDENT KINAS	1.73e+01
29	54	46.6	307	1	CD5R_MOUSE CYCLIN-DEPENDENT KINAS	1.73e+01
30	54	46.6	307	1	CD5R_MOUSE CYCLIN-DEPENDENT KINAS	1.73e+01
31	54	46.6	396	1	TCR3_ECOLI TCR3-DEPENDENT KINAS	1.73e+01
32	54	46.6	564	1	NCAP_JUNIN NUCLEOCAPSID PROTEIN (1.73e+01
33	54	46.6	754	1	YASB_SCHPO HYPOHETICAL 86.9 KD P	1.73e+01
34	54	46.6	986	1	EPH4_CHICK EPHRIN TYPE-A RECEPTOR	1.73e+01
35	54	46.6	986	1	EPH4_HUMAN EPHRIN TYPE-A RECEPTOR	1.73e+01
36	54	46.6	1278	1	DBF1_BACSU PROBABLE SERINE ACTIVA	1.73e+01
37	53	45.7	73	1	173_ASPB7 EARLY PROTEIN 173R.	2.56e+01
38	53	45.7	558	1	NCAP_LYCVU NUCLEOCAPSID PROTEIN (2.56e+01
39	53	45.7	558	1	NCAP_LYCVU NUCLEOCAPSID PROTEIN (2.56e+01
40	53	45.7	768	1	ENV_SIVAT ENVELOPE POLYPROTEIN G	2.56e+01
41	53	45.7	783	1	XPB_MOUSE DNA-REPAIR PROTEIN COM	2.56e+01
42	53	45.7	871	1	PC1_MOUSE PLASMA-CELL MEMBRANE G	2.56e+01
43	53	45.7	877	1	ENV_SIVAG ENVELOPE POLYPROTEIN G	2.56e+01
44	53	45.7	883	1	HSS2_MOUSE HEPARIN SULFATE N-DEAC	2.56e+01
45	53	45.7	1115	1	GPCR_LYMST G-PROTEIN COUPLED RECE	2.56e+01

ALIGNMENTS

RESULT	ID	T2D2_DROME	STANDARD	PRT	1213 AA.
AC	Q24325	01-NOV-1997 (REL. 35, CREATED)			
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)				
DT	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)				
DE	TRANSCRIPTION INITIATION FACTOR TFIIID 150 KD SUBUNIT (TAFII-150)				
DE	(TAFII150).				
GN	TAFI150.				
OS	DROSOPHILA MELANOGASTER (FRUIT FLY).				
OC	EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; INSECTA;				
OC	PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOPHORA; EPHYDROIDEA;				
OC	DROSOPHILINAE; DROSOPHILA.				
RN	[1]				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RC	TISSUE-EMBRYO;				
RX	MEDLINE: 94233377.				
RA	VERRIJZER C.P., YOKOMORI K., CHEN J.-L., TJIAN R.;				
RT	"Prosophiella TAFI150: similarity to yeast gene TSM-1 and specific				
RT	binding to core promoter DNA.";				
RT	SCIENCE 264:933-941(1994).				
CC	-1- FUNCTION: TAFS ARE COMPONENTS OF THE TRANSCRIPTION FACTOR IID				
CC	(TFIID) COMPLEX THAT ARE ESSENTIAL FOR MEDIATING REGULATION OF RNA				
CC	POLYMERASE TRANSCRIPTION. TAFII-150 IS AN ESSENTIAL SUBUNIT WHICH				
CC	INTERACTS DIRECTLY WITH TBP AND TAFII-250 AND BINDS TO CORE				
CC	PROMOTOR DNA.				
CC	-1- SUBUNIT: TFIIID IS COMPOSED OF TATA BINDING PROTEIN (TBP) AND A				
CC	NUMBER OF TBP-ASSOCIATED FACTORS (TAFS).				
CC	-1- SUBCELLULAR LOCATION: NUCLEAR				
CC	-1- SIMILARITY: TO YEAST TAFII-150. (TSM1).				
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation				
CC	at the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	EMBL: X79243; G54165; -				
DR	FLYBASE: FBgn0011836; Tafi150.				
DR	TRANSFAC: T02120; -				
KW	TRANSCRIPTION REGULATION; NUCLEAR PROTEIN.				
FT	DOMAIN 845 1213 BINDS TO TBP AND TAFII-250.				
FT	DOMAIN 1138 1183 HIGHLY CHARGED.				
SQ	SEQUENCE 1213 AA; 138533 MW; 2A07DC14 CRC32;				

Query Match 62.1%; Score 72; DB 1; Length 1213;
 Best Local Similarity 53.3%; Pred. No. 6.17e-03;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 DB 620 LSANDSPVLRD 634
 1 kmdptpplwkte 15

RESULT 2 STANDARD; PRT; 662 AA.
 AC P10675;
 DT 01-JUL-1989 (REL. 11, CREATED)
 DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)
 DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
 DE FASCICLIN I PRECURSOR (FAS I) (FCN).
 GN FAS1.
 OS SCHISTOCERCA AMERICANA (AMERICAN GRASSHOPPER).
 OC EURKYOTA; METAZOA; ARTROPODA; TRACHEATA; HEXAPODA; INSECTA;
 OC PTERYGOTA; ORTHOPTERA; CAELIFERA; ACRIDIDAE; CYRTACANTHACRIDINAE;
 CC SCHISTOCERCA.
 (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 88223351.
 RA ZINN K., MCALLISTER L., GOODMAN C.;
 RT "Sequence analysis and neuronal expression of fasciclin I in
 grasshopper and Drosophila.";
 RL CELL. 53:577-587(1988).
 (2)
 RP SEQUENCE OF 25-42.
 RX MEDLINE; 88276943.
 RA SNOW P.M., ZINN K., HARELSON A.L., MCALLISTER L., SCHILLING J.,
 RA BASTIANI M.J., MARK G., GOODMAN C.S.;
 RT "Characterization and cloning of fasciclin I and fasciclin II
 glycoproteins in the grasshopper.";
 RL PROC. NATL. ACAD. SCI. U.S.A. 85:5291-5295(1988).
 CC -1- FUNCTION: NEURAL CELL ADHESION MOLECULE.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
 CC -1- TISSUE SPECIFICITY: EXPRESSED ON DIFFERENT SUBSETS OF AXON BUNDLES
 (FASCICLES) IN INSECT EMBRYOS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

 DR EMBL; M20544; G160847; -
 DR PIR; A31817; A31817.
 DR A29900; A29900.
 KW CELL ADHESION; GLYCOPROTEIN; REPEAT; SIGNAL; GPI-ANCHOR.
 FT SIGNAL 1 24
 FT CHAIN 25 662
 FT REPEAT 37 165
 FT REPEAT 194 334
 FT REPEAT 339 483
 FT REPEAT 484 637
 FT CARBOHYD 54 54
 FT CARBOHYD 175 175
 FT CARBOHYD 437 437
 FT CARBOHYD 448 448
 FT CARBOHYD 488 488
 FT CARBOHYD 569 569
 SQ SEQUENCE 662 AA; 75282 MW; D5686B88 CRC32;

Query Match 55.2%; Score 64; DB 1; Length 662;
 Best Local Similarity 58.3%; Pred. No. 2.51e-01;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 DB 104 LSNDGNPPLMI 115
 1 kmdptpplwkte 15

QY 1 kmdptpplwkte 12

RESULT 3 STANDARD; PRT; 271 AA.
 AC P52197;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE THIOSULFATE SULFOTRANSFERASE (EC 2.8.1.1) (RHODANESE-LIKE PROTEIN).
 GN RHA.
 OS AZOTOBACTER VINELANDII.
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; AZOTOBACTERACEAE;
 CC AZOTOBACTER.
 (1)
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-18.
 RC STRAIN-OP / UW136;
 RX MEDLINE; 96184904.
 RA COLNAGHI R., PAGANI S., KENNEDY C., DRUMMOND M.;
 RT "Cloning, sequence analysis and overexpression of the rhodanese gene
 of Azotobacter vinelandii.";
 RL EMB. J. BIOCHEM. 236:240-248(1996).
 CC -1- CATALYTIC ACTIVITY: THIOSULFATE + CYANIDE = SULFITE + THIOCYANATE.
 CC -1- SIMILARITY: BELONGS TO THE RHODANESE FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

 DR EMBL; I43346; G1069991; -
 DR PROSITE; PS00380; RHODANESE_1; 1.
 DR PROSITE; PS00683; RHODANESE_2; 1.
 DR PFW; PF00581; Rhodanese; 2.
 KW TRANSFERASE.
 FT ACT SITE 230 230
 FT SITE 230 230
 SQ SEQUENCE 271 AA; 29629 MW; 6C39E2EE CRC32;

Query Match 53.4%; Score 62; DB 1; Length 271;
 Best Local Similarity 53.8%; Pred. No. 6.08e-01;
 Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
 DB 198 AMDPRAIRIRTD 210
 3 amdpplwkte 15

RESULT 4 STANDARD; PRT; 102 AA.
 AC P36283;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 11.4 KD PROTEIN (C4 PROTEIN).
 GN C4.
 OS TOMATO YELLOW LEAF CURL VIRUS (STRAIN AUSTRALIA) (TYLCV).
 OC VIRUSES; SSDNA VIRUSES; GEMINIVIRIDAE; BEGOMOVIRUS.
 (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 93139778.
 RA DRY I.B., RIDDEN J.E., KRAKE L.R., MULLENBUX P.M., REZAIAN M.A.;
 RT "Nucleotide sequence and genome organization of tomato leaf curl
 geminivirus.";
 RL J. GEN. VIROL. 74:147-151(1993).

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial

DR	EMBL; D90863; G1799714; -
DR	EMBL; U76961; G1684787; -

OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria
OC Primates; Catarrhini; Hominae; Homo.
RN [1]
RP Sequence from N.A.
BK Medline; 96141096.

RA MA C., STAUDT L.M.;
 RT "LAF-4 encodes a lymphoid nuclear protein with transactivation
 CC potential that is homologous to Ar-4, the gene fused to MLL in
 CC t(4;11) leukemias."
 RL BLOOD 87:734-745(1996).
 CC -1- FUNCTION: PUTATIVE TRANSCRIPTION ACTIVATOR THAT MAY FUNCTION IN
 CC LYMPHOID DEVELOPMENT AND ONCOGENESIS. BINDS, IN VITRO, TO
 CC DOUBLE-STRANDED DNA.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN LYMPHOID TISSUES,
 CC HIGHEST LEVELS BEING FOUND IN THE THYMUS.
 CC -1- SIMILARITY: TO AP4 AND OX19.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U34360; G1144493; -
 KW NUCLEAR PROTEIN; TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING.
 FT DOMAIN 413 419 POLY-SER.
 FT DOMAIN 422 432 POLY-SER.
 FT DOMAIN 440 445 POLY-SER.
 FT DOMAIN 670 679 POLY-SER.
 SQ SEQUENCE 1227 AA; 133734 MW; 5F8C8BBB CRC32;
 Query Match 50.9%; Score 59; DB 1; Length 1227;
 Best Local Similarity 40.0%; Pred. No. 2,22e+00;
 Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 Db 757 LKDSDFSLWKID 771
 Oy 1 lkamdpplwikte 15
 RESULT 8
 ID Y4VJ_RHISM STANDARD; PRT; 351 AA.
 AC Q53218;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 39.2 KD PROTEIN Y4VJ.
 GN Y4VJ.
 OS RHIZOBIUM SP. (STRAIN NGR234).
 OG PLASMID SYN PGR234A.
 CC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; RHIZOBIACEAE GROUP;
 CC RHIZOBIACEAE; RHIZOBIUM.
 CC [1]
 CC SEQUENCE FROM N.A.
 RX MEDLINE; 97305956.
 RA FREIBERG C.A., FELLAY R., BAIRBOCH A., BROUGHTON W.J., ROSENTHAL A.,
 RA PERRET X.;
 RT "Molecular basis of symbiosis between Rhizobium and legumes."
 RL NATURE 387:394-401(1997).
 RN [2]
 RN SEQUENCE OF 1-279 FROM N.A.
 RX MEDLINE; 96389014.
 RA FREIBERG C., PERRET X., BROUGHTON W.J., ROSENTHAL A.;
 RT "Sequencing the 500-kb GC-rich symbiotic replicon of Rhizobium sp.
 RT NGR34 using dye terminators and a thermostable 'sequenase': a
 RT beginning."
 RL GENOME RES. 6:590-600(1996).
 CC -1- SIMILARITY: TO ALKANAL MONOOXYGENASE ALPHA AND BETA CHAINS (EC
 CC 1.14.14.3) (BACTERIAL LUCIFERASE), BUT DISTANTLY RELATED.
 CC -1- SIMILARITY: TO Y4VJ.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z68203; E213911; -
 DR EMBL; AE000101; G2182678; -
 DR PFM; PF00296; bac_luciferase.1.
 KW HYPOTHETICAL PROTEIN; OXIDOREDUCTASE; MONOOXYGENASE; PLASMID.
 SQ SEQUENCE 351 AA; 39158 MW; CA394E88 CRC32;
 Query Match 50.0%; Score 58; DB 1; Length 351;
 Best Local Similarity 77.8%; Pred. No. 3.38e+00;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Db 167 PTPPIWIAF 175
 Oy 6 ptpplwik 14
 RESULT 9
 ID RRPL_AHSV9 STANDARD; PRT; 1305 AA.
 AC Q70695;
 DT 15-DEC-1998 (REL. 37, CREATED)
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (VPI).
 GN S1.
 OS AFRICAN HORSE SICKNESS VIRUS 9 (AHSV-9) (AFRICAN HORSE SICKNESS VIRUS
 OS (SEROTYPE 9)).
 CC VIRUSES; DSRNA VIRUSES; REOVIRIDAE; ORBIVIRUS.
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE; 98202706.
 RA VREDE F.T., HUISMANS H.;
 RT "Sequence analysis of the RNA polymerase gene of African horse
 RT sickness virus."
 RL ARCH. VIROL. 143:413-419(1998).
 CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE
 CC + RNA(N).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U94887; G2961464; -
 KW TRANSFERASE; NUCLEOTIDYLTRANSFERASE; RNA-DIRECTED RNA POLYMERASE.
 SQ SEQUENCE 1305 AA; 150293 MW; 251BD435 CRC32;
 Query Match 50.0%; Score 58; DB 1; Length 1305;
 Best Local Similarity 50.0%; Pred. No. 3.38e+00;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 Db 960 VKLIDSPPLNNEI 973
 Oy 1 lkamdpplwik 14
 RESULT 10
 ID CO4_HUMAN STANDARD; PRT; 1741 AA.
 AC P01028;
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE COMPLEMENT C4 PRECURSOR [CONTAINS: C4A ANAPHYLATOXIN].
 GN C4A AND C4B.
 OS HOMO SAPIENS (HUMAN).
 CC EUDAROTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 CC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
 RN [1]

RA SEQUENCE OF 1-22 AND 1056-1225 FROM N.A.
 RA MEDLINE; 85156269.
 RA BELT K.T., YU C.Y., CARROLL M.C., PORTER R.R.;
 RT "Polymorphism of human complement component C4.";
 RL IMMUNOGENETICS 21:173-180(1985).
 RN [2]
 RP SEQUENCE OF 20-1741 FROM N.A.
 RC TISSUE-LIVER;
 RX MEDLINE; 84156544.
 RA BELT K.T., CARROLL M.C., PORTER R.R.;
 RT "The structural basis of the multiple forms of human complement component C4.";
 RL CELL 36:907-914(1984).
 RN [3]
 RP SEQUENCE OF 680-756.
 RX MEDLINE; 81264286.
 RA MOON K.E., GORSKI J.P., HUGLI T.E.;
 RT "Complete primary structure of human C4a anaphylatoxin.";
 RL J. BIOL. CHEM. 256:8685-8692(1981).
 RN [4]
 RP SEQUENCE OF 957-1044.
 RX MEDLINE; 82182029.
 RA CAMPBELL R.D., GAGNON J., PORTER R.R.;
 RT "Amino acid sequence around the thiol and reactive acyl groups of human complement component C4.";
 RL BIOCHEM. J. 199:359-370(1981).
 RN [5]
 RP SEQUENCE OF 990-1037.
 RX MEDLINE; 82150875.
 RA HARRISON R.A., THOMAS M.L., TRICK B.F.;
 RT "Sequence determination of the thiolester site of the fourth component of human complement.";
 RL PROC. NATL. ACAD. SCI. U.S.A. 78:7388-7392(1981).
 RN [6]
 RP SEQUENCE OF 1-21 FROM N.A.
 RX MEDLINE; 94282044.
 RA SARGENT C.A., ANDERSON M.J., HSIEH S.L., KENDALL E.,
 RA GOMEZ-ESCOBAR N., CAMPBELL R.D.;
 RT "Characterisation of the novel gene G11 lying adjacent to the complement C4a gene in the human major histocompatibility complex.";
 RL HUM. MOL. GENET. 3:481-488(1994).
 RN [7]
 RP STRUCTURAL BASIS OF POLYMORPHISM.
 RX MEDLINE; 87080272.
 RA YU C.Y., BELT K.T., GILES C.M., CAMPBELL R.D., PORTER R.R.;
 RT "Structural basis of the polymorphism of human complement components C4a and C4b: gene size, reactivity and antigenicity.";
 RL EMBO J. 5:2873-2881(1986).
 RN [8]
 RP VARIANT C4A6 ALLOTYPES.
 RX MEDLINE; 92242905.
 RA ANDERSON M.J., MILLNER C.M., COTTON G.H., CAMPBELL R.D.;
 RT "The coding sequence of the hemolytically inactive C4a6 allele of human complement component C4 reveals that a single arginine to tyrosine substitution at beta-chain residue 458 is the likely cause of the defect.";
 RL J. IMMUNOL. 148:2795-2802(1992).
 RN [9]
 RP FUNCTION: C4 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE CLASSICAL PATHWAY OF THE COMPLEMENT SYSTEM. IT IS PROCESSED BY ACTIVATED C1 WHICH REMOVE FROM THE ALPHA CHAIN THE C4A ANAPHYLATOXIN.
 CC -1- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C4, C4A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND BASOPHILIC LEUKOCYTES.
 CC -1- SUBUNIT: THIS PROTEIN IS SYNTHESIZED AS A SINGLE-CHAIN PRECURSOR AND, PRIOR TO SECRETION, IS ENZYMATICALLY CLEAVED TO FORM A TRIMER OF NONIDENTICAL CHAINS (ALPHA, BETA, AND GAMMA).
 CC -1- POLYMORPHISM: HUMAN COMPLEMENT COMPONENT C4 IS POLYMORPHIC WITH AT LEAST TWO LOCI, C4A AND C4B. 13 ALLELES OF C4A AND 22 ALLELES OF C4B HAVE BEEN DETECTED. THE ALLELE SHOWN HERE IS C4A4.
 CC -1- POLYMORPHISM: THE C4A ALLELES CARRY THE BLOOD GROUP RODGERS WHILE

CC THE C4B ALLELES CARRY THE BLOOD GROUP CHIDO.
 CC DISEASE: THE C4A6 ALLOTYPES IS TOTALLY DEFICIENT IN HEMOLYTIC ACTIVITY.
 CC -1- C4A ALLOTYPES REACT MORE RAPIDLY WITH THE AMINO GROUP OF PEPTIDE ANTIGENS WHILE C4B ALLOTYPES REACT MORE RAPIDLY WITH THE HYDROXYL GROUP OF CARBOHYDRATE ANTIGENS.
 CC -1- C4 IS A MAJOR HISTOCOMPATIBILITY COMPLEX CLASS-III PROTEIN.
 CC -1- SIMILARITY: TO C3, C5 AND ALPHA-2-MACROGLOBULIN.
 CC -1- SIMILARITY: CONTAINS 1 ANAPHYLATOXIN-LIKE DOMAIN.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb.ch/announce/> or send an email to license@isb.ch).
 CC -----
 CC EMBL; M14823; G553211; -
 CC EMBL; K02403; G443671; ALT_SEQ.
 CC EMBL; M14824; G553210; -
 CC EMBL; X77491; G453411; -
 CC PIR; A01262; C4HU.
 CC PIR; A17265; A17265.
 CC PIR; A29177; A29177.
 CC PIR; B20807; B20807.
 CC SWISS-2DPAGE; P01028; HUMAN.
 CC MIM; 120790; -
 CC MIM; 120810; -
 CC MIM; 120820; -
 CC PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
 CC PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
 CC PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
 CC PRAM; P00207; A2M; 1.
 CC HSSP; P01031; IKJS.
 CC COMPLEMENT PATHWAY; PLASMA; GLYCOPROTEIN; MHC III; SIGNAL;
 CC INFLAMMATORY RESPONSE; POLYMORPHISM; DISEASE MUTATION;
 CC BLOOD GROUP ANTIGEN.
 CC SIGNAL.
 CC FT CHAIN 1 19
 CC FT PROPEP 20 675
 CC FT CHAIN 676 679
 CC FT PROPEP 680 1443
 CC FT CHAIN 1444 1450
 CC FT CHAIN 1451 1741
 CC FT PEPTIDE 680 756
 CC FT DOMAIN 702 736
 CC FT DISULFID 703 735
 CC FT DISULFID 703 735
 CC FT THIOLEST 716 736
 CC FT THIOLEST 1010 1013
 CC FT CARBOHYD 226 226
 CC FT CARBOHYD 862 862
 CC FT CARBOHYD 1328 1328
 CC FT CARBOHYD 1391 1391
 CC FT VARIANT 477 477
 CC FT VARIANT 726 726
 CC FT VARIANT 1073 1073
 CC FT VARIANT 1120 1125
 CC FT VARIANT 1176 1176
 CC FT VARIANT 1201 1201
 CC FT VARIANT 1207 1207
 CC FT VARIANT 1210 1210
 CC FT VARIANT 1286 1286
 CC FT VARIANT 1286 1286
 CC FT CONFLICT 727 727
 CC FT CONFLICT 1013 1013
 CC SO SEQUENCE 1741 AA; 192335 MF; 405A8805 CRC32;
 CC -----
 CC Query Match 50.0%; Score 58; DB 1; Length 1741;
 CC Best Local Similarity 50.0%; Pred. No. 3.38e+00;
 CC Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 CC Db 1253 PMPOAPALMTET 1264
 CC :|:|:|:|:|:|

OY 3 andptplwkt 14

RESULT 11 STANDARD: PRT; 209 AA.

AC 000927;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
 DE YOP PROTEINS TRANSLLOCATION PROTEIN K (LOW CALCIUM RESPONSE LOCUS
 DE PROTEIN K).
 GN YSCK OR LCRRB.
 OS YERSINIA PSEUDOTUBERCULOSIS.
 CC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
 CC YERSINIA.
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN-YPLII:
 CC MEDLINE; 92250432.
 CC RIMPIIAINEN M., FORSBERG A., WOLF-WATZ H.;
 CC "A novel protein, Loro, involved in the low-calcium response of
 CC Yersinia pseudotuberculosis shows extensive homology to YopH.";
 CC J. BACTERIOL. 174:3355-3363(1992)
 CC -1- FUNCTION: BELONGS TO AN OPERON INVOLVED IN THE TRANSLLOCATION OF
 CC YOP PROTEINS ACROSS THE BACTERIAL MEMBRANES OR IN THE SPECIFIC
 CC CONTROL OF THIS FUNCTION.
 CC -1- INDUCTION: AT 37 DEGREES CELSIUS IN THE ABSENCE OF CALCIUM.
 CC -1- SIMILARITY: HIGH, WITH Y ENTEROCOLITICA CORRESPONDING ORF.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL; M83986; G155467; -
 CC PLASMID; VIRULENCE.
 CC SEQUENCE 209 AA; 23991 MW; 7A24E44E CRC32;

Query Match 49.1%; Score 57; DB 1; Length 209;
 Best Local Similarity 53.8%; Pred. No. 5.12e+00;
 Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Db 152 LAEMEPOPANCK 164
 | | | | |
 | | | | |
 OY 1 lkandptplwkt 13

RESULT 12 STANDARD: PRT; 209 AA.

AC 001252;
 DT 01-APR-1993 (REL. 25, CREATED)
 DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
 DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
 DE YOP PROTEINS TRANSLLOCATION PROTEIN K.
 GN YSCK.
 OS YERSINIA ENTEROCOLITICA.
 CC PLASMID PYV.
 CC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
 CC YERSINIA.
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN-439-80 / SEROTYPE O:9;
 CC MEDLINE; 91317716.
 CC MICHIELS T., VANDOUTEGHEM J.-C., DE ROUVROIT C., CHINA B., GUSTIN A.,
 CC BOUDRY P., CORNELIS G.R.;
 CC "Analysis of yitC, an operon involved in the secretion of Yop
 CC proteins by Yersinia enterocolitica.";
 CC J. BACTERIOL. 173:4994-5009(1991).
 CC -1- FUNCTION: BELONGS TO AN OPERON INVOLVED IN THE TRANSLLOCATION OF

CC YOP PROTEINS ACROSS THE BACTERIAL MEMBRANES OR IN THE SPECIFIC
 CC CONTROL OF THIS FUNCTION.
 CC -1- INDUCTION: AT 37 DEGREES CELSIUS IN THE ABSENCE OF CALCIUM.
 CC -1- SIMILARITY: HIGH, WITH Y PSEUDOTUBERCULOSIS CORRESPONDING ORF.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL; M74011; G155560; -
 CC PIR; B40049; B40049.
 CC PLASMID; VIRULENCE.
 CC SEQUENCE 209 AA; 23998 MW; CDB29D0F CRC32;

Query Match 49.1%; Score 57; DB 1; Length 209;
 Best Local Similarity 53.8%; Pred. No. 5.12e+00;
 Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Db 152 LAEMEPOPANCK 164
 | | | | |
 | | | | |
 OY 1 lkandptplwkt 13

RESULT 13 STANDARD: PRT; 355 AA.

AC P45341;
 DT 01-NOV-1995 (REL. 32, CREATED)
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE PUTATIVE UNDECAPRENYL-PHOSPHATE ALPHA-N-ACETYLGLUCOSAMINYLTRANSFERASE
 DE (EC 2.4.1.1).
 GN RFE OR H11716.
 OS HAEMOPHILUS INFLUENZAE.
 CC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURILLACEAE;
 CC HAEMOPHILUS.
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN-RD / KW20;
 CC MEDLINE; 95350630.
 CC RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,
 CC KEELAVAGE A.R., BULL C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,
 CC KENNERLY K., SETTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,
 CC SCOTT J.D., SHIRLEY R., LIU L.-I., GLOER A., KELLEY J.M.,
 CC WEIDMAN J.E., PHILLIPS C.A., SPRIGGS T., HEDBLUM E., COTTON M.D.,
 CC UETTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,
 CC FINE L.D., FRITCHMAN J.L., FUHRMAN J.L., GEOGHAGEN N.S.M.,
 CC GNEHM C.L., MCDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
 CC VENTER J.C.;
 CC "Whole-genome random sequencing and assembly of Haemophilus
 CC influenzae Rd.";
 CC SCIENCE 269:496-512(1995).
 CC -1- FUNCTION: MAY BE THE TUNICAMYCIN SENSITIVE TRANSFERASE THAT
 CC CATALYZES THE SYNTHESIS OF GLUCNA-C-PYROPHOSPHORYLUNDECAPRENOYL
 CC (LIPID I), THE FIRST LIPID-LINKED INTERMEDIATE INVOLVED IN ECA
 CC SYNTHESIS (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: UDP-N-ACETYL-D-GLUCOSAMINE + UNDECAPRENYL
 CC PHOSPHATE -> DMP + UNDECAPRENYL N-ACETYL-ALPHA-D-GLUCOSAMINYL
 CC PYROPHOSPHATE.
 CC -1- PATHWAY: INVOLVED IN THE SYNTHESIS OF ENTEROBACTERIAL COMMON
 CC ANTIGEN (ECA) AND REQUIRED FOR SYNTHESIS OF LIPOPOLYSACCHARIDE
 CC O-SIDE CHAINS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: SOME, TO THE DOLICHYL-PHOSPHATE ALPHA-N-ACETYL-
 CC GLUCOSAMINYLTRANSFERASES OF YEAST AND MAMMALS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U32844; G1574571; -

DR TIGR: H11716; -

KW LIPOPOLYSACCHARIDE BIOSYNTHESIS; GLYCOSYLTRANSFERASE; TRANSFERASE;

KM TRANSFERASE.

FT TRANSFERASE 1 21 POTENTIAL.

FT TRANSFERASE 39 59 POTENTIAL.

FT TRANSFERASE 63 83 POTENTIAL.

FT TRANSFERASE 100 120 POTENTIAL.

FT TRANSFERASE 123 143 POTENTIAL.

FT TRANSFERASE 151 171 POTENTIAL.

FT TRANSFERASE 182 202 POTENTIAL.

FT TRANSFERASE 208 228 POTENTIAL.

FT TRANSFERASE 237 257 POTENTIAL.

FT TRANSFERASE 289 309 POTENTIAL.

FT TRANSFERASE 316 336 POTENTIAL.

FT TRANSFERASE 355 AA; 40070 MW; 4CA1FF0A CRC32;

Query Match 49.1%; Score 57; DB 1; Length 355;

Best Local Similarity 50.0%; Pred. No. 5.12e+00;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DB 236 PNPVTALMT 245

QY 3 amdpplp1w1 12

RESULT 14
 ID 14PG_RHISN STANDARD; PRT; 192 AA.
 AC P55616;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 21.1 KD PROTEIN 14PG/14SC.
 GN 14PG AND 14SC.
 OS RHIZOBIUM SP. (STRAIN NGR234).
 OC PLASMID SYM PNCGR234A.
 OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; RHIZOBIACEAE GROUP;
 OC RHIZOBIACEAE; RHIZOBIUM.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 97305956.
 RA FREIBERG C.A., FELLAY R., BAIRROCH A., BROUGHTON W.J., ROSENTHAL A.,
 RA PERRET X.;
 RL "Molecular basis of symbiosis between Rhizobium and legumes";
 RL NATURE 387:394-401(1997).
 CC -1 SIMILARITY: STRONG, TO ACETOBACTER XYLINUM IS1268 ORFA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AE000090; G2182576; -

DR EMBL: AE000095; G2182617; -

DR HYPOTHETICAL PROTEIN; PLASMID; TRANSPORTABLE ELEMENT.

SEQUENCE 192 AA; 21094 MW; DB6869B CRC32;

Query Match 48.3%; Score 56; DB 1; Length 192;

Best Local Similarity 66.7%; Pred. No. 7.73e+00;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

DB 166 SPLMKRSE 174

QY 7 ttplwkte 15

RESULT 15
 ID SRG1_CAMEL STANDARD; PRT; 328 AA.
 AC P46570;

DT 01-NOV-1995 (REL. 32, CREATED)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)

DE SRC-1 PROTEIN.

GN SRC-1 OR C18F10.4.

OS CAENORHABDITIS ELEGANS.

OC EUKARYOTA; METAZOA; NEURATA; SECRETINER; RHABDITIDA;

OC RHABDITIDA; RHABDITIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA LATREILLE P.;
 RL SUBMITTED (MAY-1994) TO EMBL/GENBANK/DDSI DATA BANKS.

CC -1 SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).

CC -1 SIMILARITY: BELONGS TO THE C.ELEGANS RECEPTOR-LIKE PROTEIN SRG

CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U00049; G465105; -

DR WORMPEP; C18F10.4; CE00781.

DR TRANSFERASE; MULTIGENE FAMILY.

FT TRANSFERASE 25 45 POTENTIAL.

FT TRANSFERASE 59 79 POTENTIAL.

FT TRANSFERASE 118 138 POTENTIAL.

FT TRANSFERASE 148 168 POTENTIAL.

FT TRANSFERASE 197 217 POTENTIAL.

FT TRANSFERASE 234 254 POTENTIAL.

FT TRANSFERASE 298 318 POTENTIAL.

FT TRANSFERASE 328 AA; 38276 MW; 24F659EA CRC32;

Query Match 48.3%; Score 56; DB 1; Length 328;

Best Local Similarity 50.0%; Pred. No. 7.73e+00;

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

DB 152 LTMATSPCLMT 163

QY 1 lkamdpplp1w1 12

Search completed: Wed Dec 29 14:06:44 1999
 Job time : 8 secs.

 RELEASE (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msearch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Wed Dec 29 14:06:03 1999; Maspar time 3.68 Seconds
 Tabular output not generated. 163.314 Million cell updates/sec

Title: >09-300612
 Description: (1-15) from baskar-09-300612.pep
 Perfect Score: 116
 Sequence: 1 lkamdpptplwikte 15

Scoring table: PAM 150
 Gap 15
 Searched: 122810 segs, 40068593 residues
 Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: plf60
 1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 26.686; Variance 40.943; scale 0.652

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description	Pred. No.
1	91	78.4	237	2	A42013	alpha-1-B-glycoprotein
2	72	62.1	1213	2	A54063	TARA-binding protein
3	65	56.0	115	2	D71194	hypothetical protein
4	64	55.2	225	2	C70045	two-component respons
5	64	55.2	662	2	A29900	fasciclin I precursor
6	62	53.4	271	2	S62187	thiosulfate sulfurtra
7	61	52.6	97	2	S59888	C4 protein - tomato
8	59	50.9	102	1	J01890	C4 protein - tomato
9	59	50.9	331	2	G65004	Div protein - Escheri
10	59	50.9	564	1	VHXPMY	major structural nucl
11	58	50.0	184	2	B20807	complement C4B - huma
12	58	50.0	609	2	S65208	probable membrane pro
13	58	50.0	1744	1	C4HU	complement C4A precu
14	57	49.1	209	2	B40049	viLC-region hypotheri
15	57	49.1	355	2	A64138	rfe protein - Haemoph
16	56	48.3	138	2	B69465	transcription regulat
17	56	48.3	162	2	C70829	hypothetical protein
18	56	48.3	324	2	S46356	pupr protein - Pseudo
19	56	48.3	517	2	S69347	steroid libeta-monoox
20	55	47.4	132	2	J02270	hypothetical 15.2K pr
21	55	47.4	263	2	S76994	hypothetical protein
22	55	47.4	319	2	UC5644	acyl-CoA thioesterase
23	55	47.4	347	2	BV0710	hypothetical protein

RESULT	ENTRY	DESCRIPTION	ALIGNMENTS
24	55	47.4	357 1 TVHUL2
25	55	47.4	425 2 C65059
26	55	47.4	446 2 B70972
27	55	47.4	453 2 S56657
28	55	47.4	522 2 S34190
29	55	47.4	552 2 A55760
30	55	47.4	594 2 E64622
31	55	47.4	594 2 B71893
32	55	47.4	680 1 XJBYRX
33	55	47.4	864 1 VGLUG4
34	55	47.4	932 2 A31898
35	54	46.6	41 2 PL0028
36	54	46.6	89 2 J00726
37	54	46.6	89 2 T00725
38	54	46.6	307 2 S50850
39	54	46.6	307 2 S50861
40	54	46.6	365 2 YTEC32
41	54	46.6	365 2 YTEC32
42	54	46.6	509 2 T00793
43	54	46.6	564 2 S06896
44	54	46.6	564 1 VHXPMY
45	54	46.6	967 2 S58360

RESULT 1
 ENTRY A42013
 TITLE alpha-1-B-glycoprotein - North American opossum (fragments)
 ORGANISM #formal_name Didelphis virginiana, Didelphis marsupialis
 #journal Virginia #common_name North American opossum
 DATE 10-Jul-1992 #sequence_revision 10-Jul-1992 #text_change 31-Dec-1993

ACCESSIONS
 REFERENCE A42013
 #authors Catanesse, J.J.; Kress, L.F.
 #journal Biochemistry (1992) 31:410-418
 #title Isolation from opossum serum of a metalloproteinase inhibitor homologous to human alpha1B-glycoprotein.
 #cross-references NID:32118834
 #accession A42013
 #status preliminary
 ##molecule-type mRNA
 ##residues 1-237 #label CAT
 ##cross-references GP:J05356
 KEYWORDS glycoprotein
 SUMMARY #length 237 #checksum 3610

Query Match 78.4%; Score 91; DB 2; Length 237;
 Best Local Similarity 80.0%; Pred. NO. 5.41e-06;
 Matches 12; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

DB 1 LKANDTTPRLMIETE 15
 ||||| -|| |||| ||
 Qy 1 lkamdpptplwikte 15

RESULT 2
 ENTRY A54063
 TITLE TARA-binding protein-associated factor II - fruit fly
 (Drosophila sp.)
 ORGANISM #formal_name Drosophila sp.
 DATE 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change 17-Mar-1999

ACCESSIONS
 REFERENCE A54063
 #authors Verrijzer, C.P.; Yokomori, K.; Chen, J.L.; Tjian, R.
 #journal Science (1994) 264:933-941
 #title Drosophila TARF-II 150: similarity to yeast gene TSM-1 and specific binding to core promoter DNA.
 #cross-references NID:94233377
 #accession A54063
 #status preliminary; not compared with conceptual translation

##molecule_type mRNA
##residues 1-1213 ##label VER
GENETICS
#gene
#FlyBase:Tafl150
#cross-references FlyBase:FBgn0011836
SUMMARY #length 1213 #molecular-weight 138533 #checksum 9849

Query Match 62.1%; Score 72; DB 2; Length 1213;
Best Local Similarity 53.3%; Pred. No. 3.07e-02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 620 ISAMDSPVIMIRD 634
| | | | | : | | | | :
QY 1 lkamdpctplwikr 15

RESULT 3
ENTRY D71194 #type complete
TITLE hypothetical protein PH1828 - *Pyrococcus horikoshii*
ORGANISM #formal_name *Pyrococcus horikoshii*
DATE 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change
D71194
A71000
Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.;
Hino, Y.; Yamamoto, S.; Sekine, M.; Baba, S.; Kosugi, H.;
Hosoyama, A.; Nagai, Y.; Sakai, M.; Ogura, K.; Otsuka, R.;
Nakazawa, H.; Takamiya, M.; Ohfuku, Y.; Funahashi, T.;
Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Oguchi,
A.; Aoki, K.; Yoshizawa, T.; Nakamura, Y.; Robb, F.T.;
Horikoshi, K.; Masuch, Y.; Shizuya, H.; Kikuchi, H.
DNA Res. (1998) 5:55-76
#journal
#title Complete sequence and gene organization of a
hyper-thermophilic archaebacterium, *Pyrococcus horikoshii*
OT3.
#cross-references M01D:98344137
#accession D71194
#status preliminary; nucleic acid sequence not shown;
translation not shown
##molecule_type DNA
##residues 1-115 ##label KAW
#cross-references GB:A000007; NID:q3236134; PID:d1031890; PID:q3258264
#experimental_source strain OT3
#note this accession replaces an interim accession for a
sequence replaced by GenBank

GENETICS
#gene
SUMMARY PH1828
#length 115 #molecular-weight 12209 #checksum 6013

Query Match 56.0%; Score 65; DB 2; Length 115;
Best Local Similarity 50.0%; Pred. No. 5.81e-01;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 88 TISPPSPIMWTS 99
| | | | | : | | | | :
QY 3 amdpctplwikr 14

RESULT 4
ENTRY C70045 #type complete
TITLE two-component response regulator [YvqB] homolog yvqA -
#formal_name *Bacillus subtilis*
ORGANISM #formal_name *Bacillus subtilis*
DATE 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
C70045
A63580
Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;
Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessières, P.;
Bollati, A.; Borcherdt, S.; Boriss, R.; Boursier, L.; Brans,
A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.;
Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
Choi, S.K.; Codani, J.J.; Conerton, I.F.; Cummings, N.J.;

Daniel, R.A.; Denliot, F.; Devine, K.M.; Duesterhoeft, A.;
Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.;
Fabret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita,
M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Galleron, N.; Ghim,
S.Y.; Glaeser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.;
Gusteppl, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood,
C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.;
Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.;
Kashihara, Y.; Kleier-Blanchard, M.; Klein, C.; Kobayashi,
I.; Koeltter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.;
Kurita, K.; Lapidus, A.; Lardinois, S.; Lauber, J.;
Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.;
Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno,
M.; Moestl, D.; Nakai, S.; Noack, M.; Noone, D.; O'Reilly,
M.; Ogawa, K.; Ogilwa, A.; Oudega, B.; Park, S.H.; Parro,
V.; Pohl, T.M.; Portetelle, D.; Porwolik, S.; Prescott,
A.M.; Prescan, E.; Fujic, P.; Punelle, B.; Rapoport, G.;
Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.;
Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, E.;
Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.;
Sekowska, A.; Seror, S.J.; Settor, P.; Shin, B.S.; Soldo,
B.; Sorokin, A.; Taccout, E.; Takagi, T.; Takahashi, H.;
Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.;
Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.;
Vandenbol, M.; Vannier, F.; Vassarotti, A.; Vieri, A.;
Wandut, R.; Wedler, E.; Wedler, H.; Weltzenegger, T.;
Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto,
K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumbstein, E.;
Yoshikawa, H.; Danchin, A.
Nature (1997) 390:249-256
#journal
#title The complete genome sequence of the Gram-positive bacterium
Bacillus subtilis.
#cross-references M01D:98044033
#accession C70045
#status preliminary; nucleic acid sequence not shown;
translation not shown
##molecule_type DNA
##residues 1-225 ##label KUN
#cross-references GB:A009120; GB:AL009126; NID:q2635613; PID:e1184380;
PID:q2635798
#experimental_source strain 168

GENETICS
#gene
CLASSIFICATION yvqA
KEYWORDS #superfamily ompr protein; response regulator homology
FEATURES
5-113
#domain response regulator homology #label RRH
#binding_site phosphate (Asp) (covalent) #status
predicted

SUMMARY #length 225 #molecular-weight 26212 #checksum 1925

Query Match 55.2%; Score 64; DB 2; Length 225;
Best Local Similarity 63.6%; Pred. No. 8.72e-01;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 40 KKMTPSPIMI 50
| | | | | : | | | | :
QY 2 kamdpctplwi 12

RESULT 5
ENTRY A29900 #type complete
TITLE fasciclin I precursor - American bird grasshopper
ORGANISM #formal_name *Schistocerca americana* #common_name American
bird grasshopper
DATE 22-Aug-1988 #sequence_revision 22-Aug-1988 #text_change
A29900
A29900
Zinn, K.; McAllister, L.; Goodman, C.S.
Cell (1988) 53:577-587
Sequence analysis and neuronal expression of fasciclin I in
grasshopper and *Drosophila*.

##cross-references MUID:88223351
#accession A29900 Preliminary
##status
##molecule_type mRNA
##residues 1-662 ##label ZIN
#cross-references GB:M20544; GB:J03787; NID:g160846; PID:g160847
SUMMARY #length 662 #molecular-weight 75282 #checksum 9667

Query Match 55.2%; Score 64; DB 2; Length 662;
Best Local Similarity 58.3%; Pred. No. 8.72e-01;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Db 104 LSEIDGNPPLMT 115
OY 1 lkamdpplwlt 12

RESULT 6
ENTRY S62187 #type complete
TITLE thiosulfate sulfurtransferase (EC 2.8.1.1) - Azotobacter vinelandii
ORGANISM rhodanese
#formal_name Azotobacter vinelandii
DATE 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change
ACCESSIONS S62187; S62211
REFERENCE S62187; S62211
#authors Colnaghi, R.; Pagani, S.; Kennedy, C.; Drummond, M.
#journal Eur. J. Biochem. (1996) 236:240-248
#title Cloning, sequence analysis and overexpression of the rhodanese gene of Azotobacter vinelandii.
#cross-references MUID:96184904
#accession S62187
##molecule_type DNA
##residues 1-271 ##label COL
#cross-references EMBL:L42346; NID:g1069990; PID:g1069991
#accession S62211
##molecule_type protein
##residues 1-18 ##label COW

GENETICS rhda
KEYWORDS sulfurtransferase
FEATURE 1-271

SUMMARY #product thiosulfate sulfurtransferase #status experimental #label MAT
#length 271 #molecular-weight 29629 #checksum 3279

Query Match 53.4%; Score 62; DB 2; Length 271;
Best Local Similarity 53.8%; Pred. No. 1.94e+00;
Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 198 AMPSRALRTRTD 210
OY 3 amdpplwlt 15

RESULT 7
ENTRY S59888 #type complete
TITLE C4 protein - tomato yellow leaf curl virus
ORGANISM #formal_name tomato yellow leaf curl virus
DATE 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change
ACCESSIONS S59888
REFERENCE S58346
#authors Hong, Y.; Harrison, B.D.
#submission submitted to the EMBL Data Library, February 1995
#description Nucleotide sequences from tomato leaf curl viruses from different countries: evidence for three geographically separate branches in evolution of the coat protein of whitefly-transmitted geminiviruses.

#accession S59888 Preliminary
##status
##molecule_type DNA

##residues 1-97 ##label HON
#cross-references EMBL:Z48182; NID:9944838; PID:9974214
CLASSIFICATION #superfamily tomato yellow leaf curl virus C4 protein
SUMMARY #length 97 #molecular-weight 11026 #checksum 3058

Query Match 52.6%; Score 61; DB 2; Length 97;
Best Local Similarity 46.7%; Pred. No. 2.89e+00;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 42 LRNVOMSRPMKTE 56
OY 1 lkamdpplwlt 15

RESULT 8
ENTRY J01890 #type complete
TITLE C4 protein - tomato yellow leaf curl virus (strain Australia)
ORGANISM #formal_name tomato yellow leaf curl virus
DATE 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change
ACCESSIONS J01890
REFERENCE J01885
#authors Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaiyan, M.A.
#journal J. Gen. Virol. (1993) 74:147-151
#title Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
#accession J01890
##status translation not shown
##molecule_type DNA
##residues 1-102 ##label DRY
#cross-references GB:S53251
CLASSIFICATION #superfamily tomato yellow leaf curl virus C4 protein
SUMMARY #length 102 #molecular-weight 11410 #checksum 1606

Query Match 50.9%; Score 59; DB 1; Length 102;
Best Local Similarity 40.0%; Pred. No. 6.29e+00;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Db 44 LRARQTSPIWRTTE 58
OY 1 lkamdpplwlt 15

RESULT 9
ENTRY G65004 #type complete
TITLE Div protein - Escherichia coli (strain K-12)
ORGANISM #formal_name Escherichia coli
DATE 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change
ACCESSIONS G65004
REFERENCE A64720
#authors Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.; Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.W.; Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao, Y.
#journal Science (1997) 277:1453-1462
#title The complete genome sequence of Escherichia coli K-12.
#cross-references MUID:97426517
#accession G65004
##status preliminary; nucleic acid sequence not shown;
##molecule_type DNA
##residues 1-331 ##label BLAT
#cross-references GB:AE000321; GB:U00096; NID:g1788659; PID:g1788661; UMGF:b2321

GENETICS #experimental_source strain K-12, substrain MG1655
KEYWORDS div
SUMMARY #length 331 #molecular-weight 36668 #checksum 690

Query Match 50.9%; Score 59; DB 2; Length 331;

Best Local Similarity 50.0%; Pred. No. 6,29e+00;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Db 306 LONMARPALMI 317

QY 1 1kamdppplw1 12

RESULT 10

ENTRY VHXPMV #type complete

TITLE major structural nucleoprotein - Machupo virus

ALTERNATE_NAMES nucleocapsid protein

ORGANISM #formal_name Machupo virus

DATE 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 05-Sep-1997

ACCESSIONS S18042

REFERENCE S18042

#authors Griffiths, C.; Wilson, S.M.; Clegg, J.C.S.

#description Submitted to the EMBL Data Library, October 1991

#relationship Submitted to the nucleocapsid gene of Machupo virus: close relationship with another South American pathogenic arenavirus, Junin.

#accession S18042

#molecule_type genomic RNA

#residues 1-564 #label GRI

#cross-references EMBL:X62616; NID:960621; PID:960622

GENETICS

#map_position segment 5

CLASSIFICATION #superfamily arenavirus major nucleoprotein

KEYWORDS nucleocapsid; nucleoprotein

SUMMARY #length 564 #molecular_weight 63299 #checksum 9818

Query Match

Best Local Similarity 46.7%; Pred. No. 6,29e+00;

Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Db 368 VKRIDPTWLDIE 382

QY 1 1kamdppplw1 15

RESULT 11

ENTRY B20807 #type fragment

TITLE complement C4b - human (fragment)

ORGANISM #formal_name Homo sapiens #common_name man

DATE 21-May-1988 #sequence_revision 21-May-1988 #text_change 29-Aug-1997

ACCESSIONS B20807

REFERENCE A90845

#authors Bell, K.T.; Carroll, M.C.; Porter, R.R.

#journal Cell (1984) 36:907-914

#title The structural basis of the multiple forms of human complement component C4.

#cross-references MUID:84156544

#accession B20807

#molecule_type mRNA

#residues 1-184 #label BEL

GENETICS

#gene GDB:C4B

#cross-references GDB:119733; OMIM:120820

CLASSIFICATION #superfamily alpha-2-macroglobulin

SUMMARY #length 184 #checksum 6928

Query Match

Best Local Similarity 50.0%; Pred. No. 9,22e+00;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 137 PMPDAPALMIEF 148

QY 3 3endpplw1 14

RESULT 12

ENTRY S65208 #type complete

TITLE probable membrane protein YPL189w - yeast (Saccharomyces cerevisiae)

ALTERNATE_NAMES hypothetical protein P2201

ORGANISM #formal_name Saccharomyces cerevisiae

DATE 10-Dec-1994 #sequence_revision 31-May-1996 #text_change 21-Nov-1997

ACCESSIONS S65208; S65201

REFERENCE S65202

#authors Rieger, M.; Mueller-Auer, S.; Schaefer, M.

#description Submitted to the Protein Sequence Database, May 1996

#accession S65208

#molecule_type DNA

#residues 1-609 #label RIE

#cross-references EMBL:Z73545; NID:91370394; PID:e246916; PID:91370395; MIPS:YPL189w

REFERENCE #experimental_source strain S288C (AB972)

S65183

#authors Benes, V.; Rechmann, S.; Newtich, U.; Voss, H.; Anseorge, W.

#description Submitted to the Protein Sequence Database, May 1996

#accession S65201

#molecule_type DNA

#residues 177-609 #label BEN

#cross-references EMBL:Z73545; MIPS:YPL189w

#experimental_source strain S288C (AB972)

GENETICS

#map_position 161

KEYWORDS

FEATURES transmembrane protein

79-95

#domain transmembrane #status predicted #label TM1

#domain transmembrane #status predicted #label TM2

#domain transmembrane #status predicted #label TM3

#domain transmembrane #status predicted #label TM4

#domain transmembrane #status predicted #label TM5

#domain transmembrane #status predicted #label TM6

#domain transmembrane #status predicted #label TM7

#domain transmembrane #status predicted #label TM8

#domain transmembrane #status predicted #label TM9

#domain transmembrane #status predicted #label TM10

#length 609 #molecular_weight 71288 #checksum 4342

SUMMARY

Query Match

Best Local Similarity 54.5%; Pred. No. 9,22e+00;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 59 LSSNSPSPPLW 69

QY 1 1kamdppplw 11

RESULT 13

ENTRY C4HD #type complete

TITLE complement C4a precursor - human

CONTAINS classical-complement-pathway C3/C5 convertase (EC 3.4.21.43)

ORGANISM #formal_name Homo sapiens #common_name man

DATE 25-Feb-1985 #sequence_revision 23-Aug-1996 #text_change 24-Oct-1997

ACCESSIONS I56095; A29177; A90845; A19311; A92337; S12866;

REFERENCE A17265; A32335; A27600; I58991; I37399; A01262; A01263

#authors Yu, C.Y.

#journal J. Immunol. (1991) 146:1057-1066

#title The complete exon-intron structure of a human complement component C4a gene. DNA sequences, polymorphism, and linkage to the 21-hydroxylase gene.

#cross-references MUID:91108039

#accession I56095

#status Preliminary; translated from GB/EMBL/DBJ

#molecule_type DNA

#residues 1-1744 #label RES

#cross-references GB:M59815; NID:g179672; PID:g179674

##note 1073-Gly, 1120-Leu, 1121-Ser, 1124-I
1207-Ala, 1210-Arg were also found

20-675, 757-1446, 1454-1744	#product C4b	#status predicted	#label C4b\

1454-1744

```

680-1446      #product complement C4 alpha chain #status predicted
680-756      #label ALP\
757-845      #product C4a anaphylatoxin #status experimental #label
957-1336      C4A\
1454-1744      #region C4b-binding protein binding\
756-757      #product C4d fragment #status experimental #label C4d\
1010-1013      #product complement C4 gamma chain #status predicted
1328          #label GAM\
              #cleavage-site Arg-Ala (complement subcomponent C1s)
              #status experimental\
              #cross-link thiolester (Cys-gln) #status experimental\
              #binding-site carbohydrate (asn) (covalent) #status
              experimental
SUMMARY      #length 1744 #molecular-weight 192860 #checksum 9431

Query Match  50.0%; Score 58; DB 1; Length 1744;
Best Local Similarity 50.0%; Pred. No. 9.22e+00;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 1253 PMPOAPALWMT 1264
QY          1 1 1 1 1 1
           3 amdpplpwikt 14

RESULT 14
ENTRY    B40049      #type complete
TITLE    virc-region hypothetical protein yscK - Yersinia
          enterocolitica plasmid pY
ORGANISM #formal name Yersinia enterocolitica
DATE     21-Feb-1992 #sequence_revision 21-Feb-1992 #text_change
          09-Sep-1997

ACCESSIONS B40049
REFERENCE  A40361
           Michiels, T.; Vanooteghem, J.C.; Lambert de Rouvroit, C.;
           China, B.; Gustin, A.; Boudry, P.; Cornelis, G.R.
           J. Bacteriol. (1991) 173:4994-5009
           Analysis of virc, an operon involved in the secretion of Yop
           proteins by Yersinia enterocolitica.
           #cross-references M01D:91317716
           #accession B40049
           #status preliminary
           #molecule_type DNA
           #residues 1-209 #label MTC
           #cross-references GB:M74011; NID:g155549; PID:g155560
GENETICS
SUMMARY    plasmid
           #length 209 #molecular-weight 23998 #checksum 7033

Query Match  49.1%; Score 57; DB 2; Length 209;
Best Local Similarity 53.8%; Pred. No. 1.35e+01;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Db 152 LAAMEPOPOACK 164
QY          1 1 1 1 1 1
           1 kameppplwikt 13

RESULT 15
ENTRY    A64138      #type complete
TITLE    rfe protein - Haemophilus influenzae (strain Rd KW20)
ORGANISM #formal name Haemophilus influenzae
DATE     18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change
          17-Oct-1997

ACCESSIONS A64138
REFERENCE  A64000
           Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.;
           Kirkness, E.F.; Kerlavage, A.R.; Bult, C.J.; Tomb, J.F.;
           Dougherty, B.A.; Merriick, J.M.; McKenney, K.; Sutton, G.;
           FitzHugh, W.; Fields, C.; Gocayne, J.D.; Scott, J.;
           Shirley, R.; Liu, L.I.; Godecke, A.; Kelley, J.M.; Weidman,
           J.F.; Phillips, C.A.; Spriggs, T.; Hedblom, E.; Cotton,
           M.D.; Usterback, T.R.; Hanna, M.C.; Nguyen, D.T.; Saudek,
           D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann,

```

```

J.L.; Geoghegan, N.S.M.; Gnehm, C.L.; McDonald, L.A.;
Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, J.C.
Science (1995) 269:496-512
Whole-genome random sequencing and assembly of Haemophilus
influenzae Rd.
#cross-references M01D:95350630
#accession A64138
#status nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-355 #label TIGR
#cross-references GB:U32844; GB:L42023; NID:g1574563; PID:g1574571;
TIGR:H11716

GENETICS
KEYWORDS   rfe
SUMMARY    #length 355 #molecular-weight 40070 #checksum 678

Query Match  49.1%; Score 57; DB 2; Length 355;
Best Local Similarity 50.0%; Pred. No. 1.35e+01;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 236 PMNPVTALWT 245
QY          1 1 1 1 1 1
           3 amdpplpwikt 12

Search completed: Wed Dec 29 14:06:19 1999
Job time : 16 secs.

```

RELEASE (TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPearch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 29 14:07:02 1999; Maspar time 5.99 Seconds
136.623 Million cell updates/sec

Tabular output not generated.

Title: >09-300612
Description: (1-15) from baskar-09-300612.pep
Perfect Score: 116
Sequence: 1 lkamdpplwlkte 15

Scoring table: PAM 150
Gap 15

Searched: 179066 segs, 54579741 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

sptrembl9
1:sp:archaea 2:sp:bacteria 3:sp:fungi 4:sp:human
5:sp:invertebrate 6:sp:mammal 7:sp:mhc 8:sp:organelle
9:sp:phage 10:sp:plant 11:sp:rodent 12:sp:unclassified
13:sp:vertebrate 14:sp:virus

Statistics: Mean 26.420; Variance 36.183; scale 0.730

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	65	56.0	115	1	059492	115AA LONG HYPOHETICA
2	65	56.0	197	11	088268	ACYL-COA HYDROLASE-LIK
3	64	55.2	216	2	032304	YIRD.
4	64	55.2	225	2	032192	YVCA PROTEIN.
5	64	55.2	836	11	P97359	TAF195.
6	64	55.2	1189	4	043604	TATA BINDING PROTEIN A
7	64	55.2	1199	4	043487	COFACTOR OF INITIATOR
8	64	55.2	1199	4	060668	TBP-ASSOCIATED FACTOR
9	64	55.2	1327	4	085070	KIAA0364.
10	61	52.6	97	14	088560	V1, V2, C1, C2, C3 AND
11	61	52.6	258	5	017984	R06B9.2 PROTEIN.
12	61	52.6	1220	13	098864	PATCHED PROTEIN.
13	59	50.9	354	5	076637	T08E11.6 PROTEIN.
14	59	50.9	453	11	055171	MITOCHONDRIAL VERT-ION
15	59	50.9	100	4	014033	COMPLEMENT COMPONENT C
16	58	50.0	381	4	014835	COMPLEMENT COMPONENT C
17	58	50.0	395	2	047207	FASG PRECURSOR.
18	58	50.0	492	5	021653	COSMID R02P2.
19	58	50.0	609	3	008929	CHROMOSOME XVI READING
20	58	50.0	766	10	064790	TIF9.21.

21	58	50.0	925	10	064789	TIF9.20.	4.73e+00
22	58	50.0	1699	4	013160	COMPLEMENT CAB PRECURS	4.73e+00
23	58	50.0	1744	4	013906	COMPLEMENT COMPONENT C	4.73e+00
24	57	49.1	97	14	041338	HYPOTHETICAL 11.2 KD P	7.23e+00
25	56	48.3	138	1	028551	TRANSCRIPTIONAL REGULA	1.10e+01
26	56	48.3	162	2	053756	HYPOTHETICAL 17.6 KD P	1.10e+01
27	56	48.3	324	2	052209	PPPI AND PUDE GENES.	1.10e+01
28	56	48.3	403	5	P91736	LM-AMID AND MM-AMID-CO	1.10e+01
29	56	48.3	873	13	098949	XYL-RELATED RECEPTOR	1.10e+01
30	56	48.3	1794	14	P87515	NONSTRUCTURAL POLYPROT	1.10e+01
31	55	47.4	132	2	044307	RIBULOSE 1,5-BISPHOSPH	1.66e+01
32	55	47.4	263	3	055923	HYPOTHETICAL 28.8 KD P	1.66e+01
33	55	47.4	319	4	014734	HIV-NEF ASSOCIATED ACY	1.66e+01
34	55	47.4	319	4	015261	THIOESTERASE II.	1.66e+01
35	55	47.4	347	2	P71844	HYPOTHETICAL 37.7 KD P	1.66e+01
36	55	47.4	419	11	055137	ACYL-COA THIOESTERASE.	1.66e+01
37	55	47.4	453	10	049392	ASPARTATE AMINOTRANSFE	1.66e+01
38	55	47.4	457	10	022618	C13D.8 PROTEIN.	1.66e+01
39	55	47.4	651	5	016242	MEMBRANE GUANYLYL CYCL	1.66e+01
40	55	47.4	792	5	018866	COSMID C55C3.	1.66e+01
41	55	47.4	1075	13	042129	GUANYLYL CYCLASE C.	1.66e+01
42	55	47.4	1395	5	010463	HYPOTHETICAL 158.6 KD	1.66e+01
43	55	47.4	1449	14	065974	ORE 1.	1.66e+01
44	55	47.4	1569	10	022997	FEP23.4 PROTEIN.	1.66e+01

ALIGNMENTS

RESULT ID	1	PRELIMINARY;	PRT;	115 AA.
AC	059492;			
DT	01-ATG-1998 (TREMBLREL. 07, CREATED)			
DT	01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)			
DT	01-JUN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)			
DE	115AA LONG HYPOHETICAL PROTEIN.			
GN	P1828.			
OS	PYROCOCUS HORIKOSHII.			
OC	ARCHAEA; EURYARCHAEOTA; THERMOCOCCALES; THERMOCOCCACEAE; PYROCOCUS.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-OT3;			
RX	MEDLINE; 98344137.			
RA	KAMARABAYASI Y., SAWADA M., HORIKAWA H., HAIRAWA Y., HINO Y.,			
RA	YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSUYAMA A., NAGAI Y.,			
RA	SAKAI M., OGURA K., OTURA R., NAKAZAWA H., IKAWATA M., OHFUKU Y.,			
RA	FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OGUCHI A.,			
RA	AKOI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,			
RA	KIKUCHI H.;			
RT	*Complete Sequence and Gene Organization of the Genome of a			
RT	Hyper-thermophilic Archaeobacterium, Pyrococcus horikoshii OT3.;			
RL	DNA RES. 5:55-76(1998).			
DR	EMBL; AF000007; D1031890; -			
SQ	SEQUENCE 115 AA; 12209 MW; 1846C3BA CRC32;			
Query Match	56.0%; Score 65; DB 1; Length 115;			
Best Local Similarity	50.0%; Pred. No. 2.14e-01;			
Matches	6; Conservative	4; Mismatches	2; Indels	0; Gaps
DB	88 TISPSPLWITS 99			
OY	3 amdpplwlkt 14			
RESULT 2				
ID	088268	PRELIMINARY;	PRT;	197 AA.
AC	088268;			
DT	01-NOV-1998 (TREMBLREL. 08, CREATED)			
DT	01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)			
DT	01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)			
DE	ACYL-COA HYDROLASE-LIKE PROTEIN (FRAGMENT).			
OS	RATIOS NORVEGICUS (RAT).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;			

OC SCITROGNATHI; MURIDAE; MORINAE; RATIUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LIVER;
 RX MEDLINE; 98369597.
 RA YAMADA J., SUGA K., FURIHATA T., KIRAHARA M., WATANABE T.,
 RT "CDNA cloning and genomic organization of peroxisome
 RT proliferator-inducible long-chain acyl-CoA hydrolase from rat liver
 RT cytosol".
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 248:608-612(1998).
 DR EMBL; AB010429; D1033502; -.
 KM HYDROLASE.
 FT NON_TER
 SQ SEQUENCE 197 AA; 21593 MW; 47BDE8C CRC32;
 Query Match 56.0%; Score 65; DB 11; Length 197;
 Best Local Similarity 54.5%; Pred. No. 2,14e-01;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 Db 122 TRAMPERP1W 132
 QY :||:| |||
 1 kmdptpplw 11
 RESULT 3
 ID 032304 PRELIMINARY; PRT; 216 AA.
 AC 032304;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE YTRD
 OS BACILLUS SUBTILIS.
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
 OC BACILLUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE; 98015415.
 RA MEDINA N., VANNIER F., ROCHE B., AUTRET S., LEVINE A., SEROR S.J.;
 RT "Sequencing of regions downstream of addA (98 degrees) and ctc (289
 RT degrees) in Bacillus subtilis".
 RL MICROBIOLOGY 143:3305-3308(1997).
 DR EMBL; Z93941; E1173561; -.
 DR PRAM; PF00072; response_reg; 1.
 DR PERM; PF00486; trans_reg_C; 1.
 SQ SEQUENCE 216 AA; 25130 MW; 26FA9B8 CRC32;
 Query Match 55.2%; Score 64; DB 2; Length 216;
 Best Local Similarity 63.6%; Pred. No. 3,37e-01;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Db 31 KKMPSPHLW1 41
 QY :||:| |||
 2 kmdptpplw 12
 RESULT 4
 ID 032192 PRELIMINARY; PRT; 225 AA.
 AC 032192;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE TYOA PROTEIN.
 GN TYOA.
 OS BACILLUS SUBTILIS.
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
 OC BACILLUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RX MEDLINE; 98044033.
 RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,
 RA AZEVEDO V., BERNERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,

RA BORRIS R., BOUERSIER L., BRANS A., BRAUN M., BRIGHELL S.C., BRON S.,
 RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,
 RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,
 RA DENICOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMERSON P.T.,
 RA ERTIAN K.D., ERINGTON J., FABRET C., FERRARI E., FOUGER D.,
 RA FRITZ C., FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N.,
 RA GIM S.Y., GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G.,
 RA GISEPPI G., GUY B.J., HAGA K., HAIECH J., HARWOOD C.R., HENAUT A.,
 RA HILBERT H., HOLSAPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L.,
 RA JORIS B., KAMATA D., KASAHARA Y., KLAER-BLANCHARD M., KLEIN C.,
 RA KOBAYASHI Y., KOETTER P., KONINGSTEIN G., KROGH S., KOMANO M.,
 RA KURITA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,
 RA LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,
 RA MEDINA N., MELLADO R.P., MIZUNO M., MOESTL D., NAKAI S., NOBACK M.,
 RA NOONE D., O'REILLY M., OGAWA K., OGURA A., ODEGA B., PARK S.H.,
 RA PARRO V., POHL T.M., PORTETELLE D., POROLLIK S., PRESCOTT A.M.,
 RA PRESCAN E., PUTIC P., PURNELLE B., RAPPOPORT G., REY M., REYNOLDS S.,
 RA RIESER M., RIVOLTA C., ROCHA E., ROCHE B., ROSE M., SADATE Y.,
 RA SATO T., SCANLAN E., SCHLEICH S., SCHROETER R., SCOPFONE F.,
 RA SEKIGUCHI J., SEKOWSKA A., SEROR S.J., SERROR P., SHIN B.S., SOLDI B.,
 RA SOROKIN A., TACCONE E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,
 RA TAKETCHI M., TAMAKOSHI A., TANAKA T., TERPSTRA F., TOGNONI A.,
 RA TOSATO V., UCHIYAMA S., VANDENBOL M., VANNIER F., VASSAROTTI A.,
 RA VIARI A., WABOBT R., WEDLER E., WEDLER H., WEITZENEGGER T.,
 RA WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUMOTO K., YATA K.,
 RA YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E., YOSHIKAWA H., DANCHIN A.;
 RT "The complete genome sequence of the gram-positive bacterium Bacillus
 RT subtilis".
 RL NATURE 390:249-256(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;
 RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RA WIPAT A., BRIGHELL C.S., GUY J.B., ROSE M., EMERSON P.T.,
 RA HARWOOD C.R.;
 RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; Z99120; E1184380; -.
 DR EMBL; A223978; E1249824; -.
 DR PRAM; PF00072; response_reg; 1.
 DR PERM; PF00486; trans_reg_C; 1.
 SQ SEQUENCE 225 AA; 26212 MW; FD11BDD CRC32;
 Query Match 55.2%; Score 64; DB 2; Length 225;
 Best Local Similarity 63.6%; Pred. No. 3,37e-01;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Db 40 KKMPSPHLW1 50
 QY :||:| |||
 2 kmdptpplw 12
 RESULT 5
 ID P97359 PRELIMINARY; PRT; 836 AA.
 AC P97359;
 DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
 DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE TAF195.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
 OC SCITROGNATHI; MURIDAE; MORINAE; MOS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA HEIX J., ZOMERDIJK J.C.M.B., RAVANPAY A., TJIAN R., GRUMPT I.;
 RL SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; Y09974; E300061; -.
 SQ SEQUENCE 836 AA; 92037 MW; A640D245 CRC32;
 Query Match 55.2%; Score 64; DB 11; Length 836;

Best Local Similarity 50.0%; Pred. No. 3.37e-01;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 612 LAMELSPTRPLW 623
QY 1 lkamdptplw 12

RESULT 6
ID 043604 PRELIMINARY; PRT; 1189 AA.

AC 043604;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE TATA BINDING PROTEIN ASSOCIATED FACTOR.
GN TAFII150.

OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.

RA SEQUENCE FROM N.A.
RA MARTINEZ E., GE H., TAO Y., YUAN C.-X., ROEDER R.G.;
RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF040701: G2827282;
SQ SEQUENCE 1189 AA; 135833 MW; 180759F2 CRC32;

Query Match 55.2%; Score 64; DB 4; Length 1189;
Best Local Similarity 50.0%; Pred. No. 3.37e-01;
Matches 8; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

Db 606 LSAMDADSPFLWTRID 621
QY 1 lkamdptplw 15

RESULT 7
ID 043487 PRELIMINARY; PRT; 1199 AA.

AC 043487;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE CORPCTOR OF INITIATOR FUNCTION.
GN CIFI150.

OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.

RA SEQUENCE FROM N.A.
RA KAUFMANN J., AHRENS K., KOOP R., SMALE S.T., MULLER R.;
RT "CIFI150, a human cofactor for transcription factor IID-dependent
RT initiator function."
RL MOL. CELL. BIOL. 18:233-239(1998).
DR EMBL; AF026445: G2739087;
SQ SEQUENCE 1199 AA; 136993 MW; A6363760 CRC32;

Query Match 55.2%; Score 64; DB 4; Length 1199;
Best Local Similarity 50.0%; Pred. No. 3.37e-01;
Matches 8; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

Db 616 LSAMDADSPFLWTRID 631
QY 1 lkamdptplw 15

RESULT 8
ID 060668 PRELIMINARY; PRT; 1199 AA.

AC 060668;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE TBP ASSOCIATED FACTOR TAFII150.
GN TAFII150.

OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.

RA SEQUENCE FROM N.A.
RA GUERIN M., ROEDER R.G.R.;
RL SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF057694: G3046692;
SQ SEQUENCE 1199 AA; 136985 MW; 98118DAB CRC32;

Query Match 55.2%; Score 64; DB 4; Length 1199;
Best Local Similarity 50.0%; Pred. No. 3.37e-01;
Matches 8; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

Db 616 LSAMDADSPFLWTRID 631
QY 1 lkamdptplw 15

RESULT 9
ID 015070 PRELIMINARY; PRT; 1327 AA.

AC 015070;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE KIAA0364.
GN KIAA0364.

OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.

RA SEQUENCE FROM N.A.
RA TISSUE-BRAIN;
RX MEDLINE: 97349984.
RA NAGASE T., ISHIKAWA K., NAKAJIMA D., OHIRA M., SEKI N., MIYAJIMA N.,
RA TANAKA A., KOTANI H., NOKURA N., OHARA O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro."
RL DNA RES. 4:141-150(1997).

DR EMBL; AB002362; D1021660;
DR PFM; PF00047; I9; 10.
SQ SEQUENCE 1327 AA; 147971 MW; 0A68A862 CRC32;

Query Match 55.2%; Score 64; DB 4; Length 1327;
Best Local Similarity 58.3%; Pred. No. 3.37e-01;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 25 MDPOPELWIESN 36
QY 4 mdptplw 15

RESULT 10
ID 088560 PRELIMINARY; PRT; 97 AA.

AC 088560;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE V1, V2, C1, C2, C3 AND C4 GENES.
GN C4.

OS TOMATO YELLOW LEAF CURL VIRUS (TYLCV).
OC VIRUSES; SS DNA VIRUSES; GEMINIVIRIDAE; BEGOMOVIRUS.
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-INDIAN TOMATO LEAF CURL VIRUS (ITLVCV);
RX MEDLINE: 95363424.
RA HONG Y., HARRISON B.D.;
RT "Nucleotide sequences from tomato leaf curl viruses from different
RT countries: evidence for three geographically separate branches in
RT evolution of the coat protein of whitefly-transmitted
RT geminiviruses."
RL J. GEN. VIROL. 76:2043-2049(1995).

DR EMBL: 248182; G974214; -
SQ SEQUENCE 97 AA; 11026 MW; 62CE9CD2 CRC32;

Query Match 52.6%; Score 61; DB 14; Length 97;
Best Local Similarity 46.7%; Pred. No. 1.29e+00;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

DB 42 LRAYOMSRPMKTE 56
1 kmdptplwlt 15

RESULT 11
ID 017984 PRELIMINARY; PRT: 258 AA.
AC 017984;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
DE R0689.2 PROTEIN.
GN R0689.2.
OS CAENORHABDITIS ELEGANS.
OC EURARCTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RA BAYNES C.;
RL SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE; 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SPALDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans.";
RL NATURE 368:32-38(1994).
DR EMBL: 283237; E1348576; -
SQ SEQUENCE 258 AA; 29059 MW; CFF322C5 CRC32;

Query Match 52.6%; Score 61; DB 5; Length 258;
Best Local Similarity 50.0%; Pred. No. 1.29e+00;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

DB 79 KHLDAVPLMKRTD 92
2 kmdptplwlt 15

RESULT 12
ID 098864 PRELIMINARY; PRT: 1220 AA.
AC 098864;
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT 01-FEB-1997 (TREMBLREL. 02, LAST ANNOTATION UPDATE)
DE PATCHED PROTEIN.
GN PTC1.
OS BRACHIDANTO RERIO (ZEBRAFISH) (ZEBRA DANIO).
OC EURARCTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
OC TELEOSTEI; EUTELOSTEI; OSTARIOPHYSI; CYPRINIFORMES; CYPRINOIDAE;
OC CYPRINIDAE; RASBORINAE; DANIO.
RN [1]
RP SEQUENCE FROM N.A.
RA CONCORDET J.P., LEWIS K.E., MOORE J.W., GOODRICH L.V., JOHNSON R.L.,
RA SCOTT M.P., INGHAM P.W.;
RL DEVELOPMENT 12:2835-2846(1996).
DR EMBL: X98883; E254399; -

SQ SEQUENCE 1220 AA; 135543 MW; E18E0F10 CRC32;

Query Match 52.6%; Score 61; DB 13; Length 1220;
Best Local Similarity 46.2%; Pred. No. 1.29e+00;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

DB 53 KAVGOKAPLWIRA 65
2 kmdptplwlt 14

RESULT 13
ID 076637 PRELIMINARY; PRT: 354 AA.
AC 076637;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE T08E11.6 PROTEIN.
GN T08E11.6.
OS CAENORHABDITIS ELEGANS.
OC EURARCTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RA MEDLINE; 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SPALDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans.";
RL NATURE 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RA JOHNSON D., WANSLEY P., KRAMER J., ELLIOTT G.;
RT "The sequence of C. elegans cosmid T08E11.";
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RA WATERSTON R.;
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF077546; G3319478; -
SQ SEQUENCE 354 AA; 40564 MW; 21878A8C CRC32;

Query Match 50.9%; Score 59; DB 5; Length 354;
Best Local Similarity 50.0%; Pred. No. 3.08e+00;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

DB 45 LRHNDPTQFWF 56
1 kmdptplwlt 12

RESULT 14
ID 055171 PRELIMINARY; PRT: 453 AA.
AC 055171;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE MITOCHONDRIAL VERY-LONG-CHAIN ACYL-COA THIOESTERASE.
OS RATTUS NORVEGICUS (RAT).
OC EURARCTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCIURGNATHI; MORIDAE; MORINAE; RATTUS.
RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-WISTAR;
 RX MEDLINE; 98109736.
 RA SVENSSON, ENGBERG, AOYAMA, USUDA, ALEXSON, HASHIMOTO;
 RT "Molecular cloning and characterization of a mitochondrial peroxisome
 RL BIOCHEM. J. 329:601-608(1998).
 DR EMBL; Y09333; E307161;
 SQ SEQUENCE 453 AA; 49701 MW; CFAE7D83 CRC32;

Query Match

50.9%; Score 59; DB 11; Length 453;

Best Local Similarity 66.7%; Pred. No. 3.08e+00;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 124 AMEPERPLW 132

OY 3 amdpplw 11

RESULT 15

ID 014033 PRELIMINARY; PRT; 100 AA.

AC 014033;

DT 01-NOV-1996 (TREMBLREL. 01, CREATED)

DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)

DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE COMPLEMENT COMPONENT C4.

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;

OC CATARRHINI; HOMINIDAE; HOMO.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 83117835.

RA CARROLL M.C.; PORTER R.R.;

RT "Cloning of a human complement component C4 gene."

RL PROC. NATL. ACAD. SCI. U.S.A. 80:264-267(1983).

DR EMBL; Y00502; E3882;

DR PFM; P00207; A2M; 1.

SQ SEQUENCE 100 AA; 10629 MW; 4589044E CRC32;

Query Match

50.0%; Score 58; DB 4; Length 100;

Best Local Similarity 50.0%; Pred. No. 4.73e+00;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 59 PMPQAPALMIET 70

OY 3 amdpplwikt 14

Search completed: Wed Dec 29 14:07:20 1999
 Job time : 18 secs.

 MUSE RELEASE

 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msrch_dp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 29 14:05:18 1999; MasPar time 3.98 Seconds

Tabular output not generated. 80.053 Million cell updates/sec

Title: >09-300612

Description: (1-15) from baskar-09-300612.pap

Perfect Score: 116

Sequence: 1 lkamdpplwlkte 15

Scoring table:

PAM 150

Gap 15

Searched: 170751 seqs, 21266608 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

a-geneseq35

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29 30:part30 31:part31 32:part32 33:part33
 34:part34 35:part35 36:part36 37:part37 38:part38
 39:part39

Statistics: Mean 19.338; Variance 66.724; scale 0.290

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	116	100.0	15 30	W53841	N-terminus of opossum	8.51e-05
2	116	100.0	15 20	W11575	N-terminal peptide fr	8.51e-05
3	73	62.9	10 30	W53843	N-terminus of opossum	2.85e+00
4	72	62.1	1189 11	R56496	TATA-binding protein-	3.58e+00
5	72	62.1	1213 23	W25029	TATA-binding protein-	3.58e+00
6	72	62.1	1213 20	W06086	Drosophila TATA-bind	3.58e+00
7	60	51.7	222 34	W29670	Homo sapiens clone AM	5.27e+01
8	56	48.3	97 32	W56498	Tobacco leaf curling	1.25e+02
9	55	47.4	311 32	W58530	Human peroxisomal thi	1.35e+02
10	55	47.4	446 21	W20095	H. pylori cytoplasmic	1.55e+02
11	55	47.4	551 13	R71376	Human cystathionine b	1.55e+02
12	55	47.4	594 22	W21020	H. pylori cytoplasmic	1.55e+02
13	55	47.4	769 2	R04197	Env gene of simian im	1.55e+02
14	54	46.6	233 23	W23985	Human serine protease	1.91e+02
15	54	46.6	396 30	W56695	Tetracycline resistanc	1.91e+02
16	54	46.6	544 35	W71203	Protein encoded by OR	1.91e+02

17	54	46.6	647 38	W73376	Human HPDVB78 protein	1.91e+02
18	54	46.6	986 14	R85936	Protein tyrosine-kin	1.91e+02
19	54	46.6	986 15	R85091	EPR-like receptor pro	1.91e+02
20	54	46.6	1891 30	W52844	Amycolatopsis mediet	1.91e+02
21	54	46.6	3413 30	W52849	A. mediterranei rifam	1.91e+02
22	53	45.7	1294 36	W5798	Arabidopsis ethylene	2.36e+02
23	53	45.7	1321 17	R88423	Arabidopsis thaliana	2.36e+02
24	53	45.7	1622 32	W60163	Arabidopsis thaliana	2.36e+02
25	53	45.7	4545 26	W22611	Hybrid smgG/tylg OR1	2.36e+02
26	53	45.7	4550 26	W23716	Platenolide synthase	2.36e+02
27	53	45.7	4550 26	W23716	Platenolide synthase	2.36e+02
28	52	44.8	1434 30	W52199	Mouse patched (ptc) p	2.90e+02
29	52	44.8	1434 17	R94380	Mouse patched protein	2.90e+02
30	52	44.8	1434 37	W72968	Mouse patched protein	2.90e+02
31	52	44.8	1447 30	W52200	Human patched protein	2.90e+02
32	52	44.8	1447 37	W72969	Human patched (ptc) p	2.90e+02
33	52	44.8	1447 17	R75375	Human patched protein	2.90e+02
34	52	44.8	1618 11	R60127	Human nestin protein	2.90e+02
35	52	44.8	1618 5	R27205	Human nestin.	2.90e+02
36	52	44.8	1621 32	W60162	Arabidopsis thaliana	2.90e+02
37	52	44.8	1688 30	W52848	A. mediterranei rifam	2.90e+02
38	52	44.8	2986 8	R44432	eryA region polypepti	2.90e+02
39	52	44.8	3164 16	R94345	Hepatitis GB virus (H	2.90e+02
40	51	44.0	161 29	W55392	H. pylori ORF 14cpl01	3.57e+02
41	51	44.0	184 29	W55505	H. pylori ORF 04cpl12	3.57e+02
42	51	44.0	368 8	R42448	Lignin peroxidase.	3.57e+02
43	51	44.0	3724 26	W23718	Platenolide synthase	3.57e+02
44	51	44.0	3724 26	W22608	Platenolide synthase	3.57e+02
45	50	43.1	259 29	W53482	H. pylori ORF 13a107	4.39e+02

ALIGNMENTS

RESULT	ID	Score	Query Match	Length	DB ID	Description	Pred. No.
1	W53841	100.0%	Score 116; DB 30; Length 15;				
2	W53841	100.0%	Score 116; DB 30; Length 15;				
3	W53841	100.0%	Score 116; DB 30; Length 15;				
4	W53841	100.0%	Score 116; DB 30; Length 15;				
5	W53841	100.0%	Score 116; DB 30; Length 15;				
6	W53841	100.0%	Score 116; DB 30; Length 15;				
7	W53841	100.0%	Score 116; DB 30; Length 15;				
8	W53841	100.0%	Score 116; DB 30; Length 15;				
9	W53841	100.0%	Score 116; DB 30; Length 15;				
10	W53841	100.0%	Score 116; DB 30; Length 15;				
11	W53841	100.0%	Score 116; DB 30; Length 15;				
12	W53841	100.0%	Score 116; DB 30; Length 15;				
13	W53841	100.0%	Score 116; DB 30; Length 15;				
14	W53841	100.0%	Score 116; DB 30; Length 15;				
15	W53841	100.0%	Score 116; DB 30; Length 15;				
16	W53841	100.0%	Score 116; DB 30; Length 15;				
17	W53841	100.0%	Score 116; DB 30; Length 15;				
18	W53841	100.0%	Score 116; DB 30; Length 15;				
19	W53841	100.0%	Score 116; DB 30; Length 15;				
20	W53841	100.0%	Score 116; DB 30; Length 15;				
21	W53841	100.0%	Score 116; DB 30; Length 15;				
22	W53841	100.0%	Score 116; DB 30; Length 15;				
23	W53841	100.0%	Score 116; DB 30; Length 15;				
24	W53841	100.0%	Score 116; DB 30; Length 15;				
25	W53841	100.0%	Score 116; DB 30; Length 15;				
26	W53841	100.0%	Score 116; DB 30; Length 15;				
27	W53841	100.0%	Score 116; DB 30; Length 15;				
28	W53841	100.0%	Score 116; DB 30; Length 15;				
29	W53841	100.0%	Score 116; DB 30; Length 15;				
30	W53841	100.0%	Score 116; DB 30; Length 15;				
31	W53841	100.0%	Score 116; DB 30; Length 15;				
32	W53841	100.0%	Score 116; DB 30; Length 15;				
33	W53841	100.0%	Score 116; DB 30; Length 15;				
34	W53841	100.0%	Score 116; DB 30; Length 15;				
35	W53841	100.0%	Score 116; DB 30; Length 15;				
36	W53841	100.0%	Score 116; DB 30; Length 15;				
37	W53841	100.0%	Score 116; DB 30; Length 15;				
38	W53841	100.0%	Score 116; DB 30; Length 15;				
39	W53841	100.0%	Score 116; DB 30; Length 15;				
40	W53841	100.0%	Score 116; DB 30; Length 15;				
41	W53841	100.0%	Score 116; DB 30; Length 15;				
42	W53841	100.0%	Score 116; DB 30; Length 15;				
43	W53841	100.0%	Score 116; DB 30; Length 15;				
44	W53841	100.0%	Score 116; DB 30; Length 15;				
45	W53841	100.0%	Score 116; DB 30; Length 15;				

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 lkamdpplwtkte 15
OY 1 lkamdpplwtkte 15

RESULT 2

ID W11575 standard; peptide; 15 AA.

AC W11575;

DT 20-MAR-1997 (first entry)

DE N-terminal peptide from lethal toxin neutralising factor.

KW lethal toxin neutralising factor; LTNF; opossum; bee toxin;

KM scorpion toxin; plant toxin; bacterial toxin; venom; sting;

KV snake bite.

OS Didelphis virginiana.

PN US517692.1

PD 19-NOV-1996.

PF 10-MAY-1993; 058387.

PR 10-MAY-1993; US-058387.

PS 22-SEP-1994; US-310340.

PA (LIPF/) LIPPS B V.

PI (LIPF/) LIPPS F W.

PT LIPPS BV, LIPPS FW;

DR WPI: 97-011287/01.

FT Treatment of victims of bee or scorpion stings or plant or bacterial

FT toxins - by admin. of lethal toxin-neutralising factor or its

PT N-terminal peptide

PS Claim 7; Column 9; 9PP; English.

CC The present sequence is from the N-terminus of a 68 kD protein

CC purified from the serum of the opossum Didelphis virginiana. The

CC full-length protein is a lethal toxin neutralising factor (LTNF).

CC The use of purified LTNF or of the chemically synthesised 15mer

CC N-terminal peptide for treating victims of bee stings, scorpion

CC stings and bacterial or plant toxins is claimed. The patent

CC disclosure does not provide any evidence for neutralising activity

CC against these various toxins. There is evidence of significant

CC neutralising activity of the opossum LTNF and the 15mer peptide

CC against venom from snakes of the families Crotalidae, Elapidae,

CC Hydroiidae and Viperidae.

CC Sequence 15 AA;

SQ

Query Match 100.0%; Score 116; DB 20; Length 15;

Best Local Similarity 100.0%; Pred. No. 8.51e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 lkamdpplwtkte 15

OY 1 lkamdpplwtkte 15

RESULT 3

ID W53843 standard; peptide; 10 AA.

AC W53843;

DT 08-JUL-1998 (first entry)

DE N-terminus of opossum LTNF.

KW LTNF; lethal toxin neutralising factor; Opossum; envenomation; therapy;

KM anti-haemorrhagic protein; Elapidae; Viperidae; sea snake; snake bite;

KV sepsis; allergy; bee sting; scorpion sting; plant toxin; bacterial toxin;

KW histamine reaction treatment.

OS Didelphis virginiana.

PN US574449.1

PD 28-APR-1998.

PF 03-JUN-1996; 657163.

PR 10-MAY-1993; US-657163.

PS 22-SEP-1994; US-310340.

PA (LIPF/) LIPPS B V.

PI (LIPF/) LIPPS F W.

PT LIPPS BV, LIPPS FW;

DR WPI: 98-271108/24.

FT Lethal Toxin Neutralising Factor peptide from opossum - can

FT neutralise venom(s) from all major families of poisonous snakes

PS Claim 7; Column 11; 11PP; English.

CC This sequence represents the peptide of the invention. It is a lethal

CC Toxin Neutralising Factor (LTNF) moiety from a 68 kDa anti-haemorrhagic

CC protein derived from an opossum. The peptide can be used in a method for

CC treating a victim of envenomation from a poisonous snake, preferably a

CC poisonous snake from the family of Elapidae, Viperidae or sea snake.

CC It is useful for the treatment of snake bites, sepsis, allergies caused

CC by the environment and treatment of bee or scorpion stings or toxicities

CC caused by plant or bacterial toxins. The peptide can also be used in

CC histamine reaction treatment. The peptide can be used in envenomation

CC treatment for a variety of snakes without prior identification of the

CC snake. Being short it can be synthetically prepared rather than the

CC current production in horses, where some people can show hypersensitivity

CC to horse proteins.

CC Sequence 10 AA;

SQ

Query Match 62.9%; Score 73; DB 30; Length 10;

Best Local Similarity 100.0%; Pred. No. 2.85e+00;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 lkamdppl 10

OY 1 lkamdppl 10

RESULT 4

ID R56496 standard; Protein; 1189 AA.

AC R56496;

DT 23-MAR-1995 (first entry)

DE TARA-binding protein associated factor dTAFF150.

KW TARA-binding protein associated factor; dTAFF150; screening;

KM diagnostic; therapeutic; gene transcription regulation.

OS Drosophila.

FT key location/Qualifiers

FT misc_difference 923

FT /note- "Val or Leu"

FT misc_difference 1106

FT /note- "Arg, Pro or His"

FT misc_difference 1172

FT /note- "STOP"

FT misc_difference 1176

FT /note- "STOP"

FT W09417087-A.

PN 04-AUG-1994.

PD 28-JAN-1994; U01114.

PF 28-JAN-1994; US-013412.

PR 30-JUN-1993; US-087119.

PS 30-JUN-1993; US-087119.

PA (REGC) UNIV CALIFORNIA.

PI Comal L, Dynlacht BD, Hoey T, Ruppert S, Tanese N;

PI Tjian R, Wang E, Weinzierl ROJ;

DR WPI: 94-264019/32.

DR P-PSDB: 970733.

PT TARA-binding protein associated protein factors - and

PT corresponding nucleotide sequence and deriv. antibodies, useful

PT in screening, diagnostics and therapeutics

PS Disclosure; Page 156; 180PP; English.

CC The TARA-binding protein associated factor hTAFF150 (including

CC specific antibodies and fusion products) are used in drug screening,

CC diagnostics and therapeutics. They are used in the development of

CC specific biochemical assays for screening compounds that agonise or

CC antagonise selected transcription factors involved in regulating

CC gene expression associated with human pathology.

CC Sequence 1189 AA;

SQ

Query Match 62.1%; Score 72; DB 11; Length 1189;

Best Local Similarity 53.3%; Pred. No. 3.58e+00;

Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 589 lsamdpplwtkte 603

OY 1 lkamdpplwtkte 15

RESULT 5
ID W25029 standard; Protein: 1213 AA.
AC W25029;
DE 08-OCT-1997 (first entry)
DE TATA-binding protein associated factor, dTAFlII50.
KW TATA-binding protein associated factor; TAF; nuclear protein;
KW RNA polymerase transcription; TATA-binding protein; TBP;
KW Initiation.
OS Drosophila sp.
PN US5637686-A.
PD 10-JUN-1997.
PF 28-JAN-1993; 013412.
PR 28-JAN-1993; US-013412.
PR 30-JUN-1993; US-087119.
PR 28-JAN-1994; US-188582.
PR 09-MAY-1996; US-646715.
PA (REGC) UNIV CALIFORNIA.
PI Comai L, Dynlacht BD, Hoey T, Ruppert S, Tanese N;
PI Tjian R, Wang E, Weinzierl ROJ;
PI WPI: 97-319113/29.
DR N-PSDB; T79605.
PT Nucleic acids encoding human TATA-binding protein associated factor (TAF) peptide(s) - for production of recombinant peptide(s), used for modulating transcription of TAFs.
PS Example 1: Column 131-138; 86pp; English.
CC W25029 represents TATA-binding protein associated factor (TAF) polypeptide, dTAFlII50 (mol. weight 150kD). TAF peptides derived from dTAFlII30 alpha, dTAFlII30 beta, dTAFlII40, dTAFlII60, dTAFlII80, dTAFlII10, dTAFlII50, and dTAFlII250, their human equivalents and nucleic acids encoding them, are used to modulate transcription, including transcription initiation. TAFs are nuclear proteins involved in RNA polymerase I, II and III transcription. The peptides act by binding to a different TAF, an activator, or TBP (TATA-binding protein) or competitively inhibiting association of a TAF domain with another compound, typically a protein like TBP or another TAF, an activator, or DNA.
CC Sequence 1213 AA:
SQ

Query Match 62.1%; Score 72; DB 23; Length 1213;
Best Local Similarity 53.3%; Pred. No. 3.58e+00;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 620 lsamdsplv1w1d 634
OY 1 lkamdpplw1wkte 15

RESULT 6
ID W06086 standard; Protein: 1213 AA.
AC W06086;
DE 27-JAN-1997 (first entry)
DE Drosophila TATA-binding protein associated factor dTAFlII50 protein.
KW Drosophila TATA-binding protein; TBP associated factor; TFIID;
KW RNA polymerase II; transcription; messenger RNA; nuclear fraction;
KW holonzyme; lambda-gt10; expression library.
OS Drosophila melanogaster.
PN US534410-A.
PD 09-JUL-1996.
PF 28-JAN-1993; 013412.
PR 28-JAN-1993; US-013412.
PR 30-JUN-1993; US-087119.
PR 28-JAN-1994; US-188582.
PA (REGC) UNIV CALIFORNIA.
PI Comai L, Dynlacht BD, Hoey T, Ruppert S, Tanese N;
PI Tjian R, Wang E, Weinzierl ROJ;
PI WPI: 96-333245/33.
DR N-PSDB; T42219.
PT Screen for cpds. that bind human TATA-binding protein associated factor - by testing ability to bind to polypeptide fragments of the factor, useful as (ant)agonists of transcription factors involved in disease.
PS Examples: Column 123-132; 86pp; English.
CC This is the amino acid sequence of the Drosophila TATA-binding protein

CC (TBP) associated factor (TAF) designated TAFII60. The protein is a component of the TFIID fraction required for reconstituting RNA polymerase II in vitro transcription activity. The encoded protein has an estimated mol. wt. of 60 kD by SDS-PAGE.
CC The invention relates to purified proteins involved in transcription by RNA polymerase II, the RNA polymerase which transcribes messenger RNA. RNA polymerase II transcription proceeds in vitro upon addition of several nuclear fractions designated TFIIA, B, D, E, F, H, I and J to RNA polymerase II holoenzyme. Fraction TFIID has been shown to contain a TBP and other TAFs. Purification of TFIID and separation of its components reveals 7 proteins ranging in size from 30-250 kD.
CC Serum raised against the TFIID fraction allowed cloning of the corresp. genes from lambda-gt10 expression libraries.
CC Sequence 1213 AA:
SQ

Query Match 62.1%; Score 72; DB 20; Length 1213;
Best Local Similarity 53.3%; Pred. No. 3.58e+00;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 620 lsamdsplv1w1d 634
OY 1 lkamdpplw1wkte 15

RESULT 7
ID W29670 standard; Protein: 222 AA.
AC W29670;
DE 09-NOV-1998 (first entry)
DE Homo sapiens clone AW42_3 secreted protein.
KW Clone; secreted protein.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 2..14 /note="signal peptide"
FT W09832853-A2.
PN 30-JUL-1998.
PD 23-JAN-1998; U01396.
PR 24-JAN-1997; US-788789.
PA (GENV) GENETICS INSR INC.
PI Agostino KM, Jacobs K, Lavallie ER, Mccoy JM, Werberg D,
PI Racle LA, Spaulding V, Treacy M;
PI WPI: 98-427949/36.
DR N-PSDB; V40540.
PT New isolated polynucleotide(s) and secreted proteins - isolated from human foetal kidney, adult brain, adult salivary gland, foetal brain and adult testes cDNA libraries
PS Claim 15; Page 65-66; 109pp; English.
CC The sequence is that of a secreted protein. Such a protein can have biological activities, e.g. nutritional activity, cytokine and cell proliferation/differentiation activity, immune stimulating or suppressing activity, haematopoiesis regulating activity, tissue growth activity, activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, receptor/ligand activity, anti-inflammatory activity, cadherin/tumour CC invasion suppressor activity, tumour inhibition activity, and other activities.
CC Sequence 222 AA:
SQ

Query Match 51.7%; Score 60; DB 34; Length 222;
Best Local Similarity 60.0%; Pred. No. 5.27e+01;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 135 pmhdsplv1w1 144
OY 3 amdpplw1wkte 12

RESULT 8
ID W56498 standard; Protein: 97 AA.
AC W56498;
DE 03-SEP-1998 (revised)
DE 11-AUG-1998 (first entry)
DE Tobacco leaf curling virus gene protein C4.

KM Tobacco leaf curling virus gene; TlCV; promoter; C4 protein.
OS Tobacco leaf curling virus.
PN 310070982-A.
PD 17-MAR-1998.
PF 30-AUG-1996; 230394.
PR 30-AUG-1996; JP-230394.
PA (NORO) NORINSUISANSO KYUSHU NOGO SHIKENYO.
DR WPI; 98-233830/21.
PT Tobacco leaf curling virus gene - useful for inserting into vectors
PS for expression in, e.g. tomato plants
DE Example 4; Fig 5; 9pp; Japanese.
CC This sequence represents the C4 protein encoded by the tobacco leaf
CC curling virus (TlCV) gene of the invention. TlCV gene or its promoter can
CC be inserted into a vector for expression in plants, e.g. tobacco and
CC tomato. This sequence is believed to be encoded by the TlCV gene shown in
CC v9761.
SQ Revised entry submitted to correct the Patent Number Field
Sequence 97 AA;

Query Match 48.3%; Score 56; DB 32; Length 97;
Best Local Similarity 53.3%; Pred. No. 1.25e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

DB 42 lkarmssptwkte 56
||| : :|||
QY 1 lkamdpplwkt 15

RESULT 9
ID W58530 standard; Protein; 311 AA.
AC W58530;
DT 01-SEP-1998 (first entry)
DE Human peroxisomal thioesterase.
KW Human; peroxisomal thioesterase; PXT; diagnosis; inflammatory disease;
KM fatty acid metabolism disorder; cancer; drug screening.
OS Homo sapiens.
PN US576753-A.
PD 07-JUL-1998.
PF 11-JUN-1997; 872784.
PR 11-JUN-1997; US-872784.
PA (INCY-) INCYTE PHARM INC.
PI Corley NC, Hillman JL, Shah P;
DR WPI; 98-398039/34.
DR N-PSDB; V31109.
PT New nucleic acid encoding human peroxisomal thioesterase - useful
PT for diagnosis, treatment and prevention of fatty acid metabolism
PT disorders, cancer and inflammatory disease, and in drug screening
PS Claim 9; Fig 1; 30pp; English.
CC The present sequence represents human peroxisomal thioesterase (PXT).
CC The CDNA encoding PXT was isolated from a human BRAINOT09 clone 2150905
CC CDNA library. PXT and its agonists are used in treatment, diagnosis
CC and prevention of disorders of fatty acid metabolism e.g. neuronal
CC disorders, Zellweger or Down's syndrome, Refsum's disease, epilepsy,
CC Alzheimer's disease, schizophrenia, tardive dyskinesia, multiple sclerosis,
CC PXT DNA can also be used for diagnosis and monitoring (in hybridisation
CC or amplification methods), including determining a predisposition to
CC disease, for gene mapping and to detect PXT-encoding nucleic acid,
CC particularly after amplification. PXT can be used to raise antibodies
CC (Ab) used therapeutically, in competitive drug screens, as immunoassay
CC reagents for diagnosis and monitoring of diseases and for purification of
CC natural PXT, and to screen for specific binding agents antagonists e.g.
CC specific Ab, antisense sequences or ribozymes which are used in treatment
CC of cancers and inflammatory diseases (e.g. adult respiratory distress
CC syndrome, asthma, diabetes mellitus, microbial and other infections).
CC Administration of proteins, agonist or antagonists, optionally combined
CC with conventional pharmaceuticals providing a synergistic effect, is
CC e.g. orally, by injection or topically.
SQ Sequence 311 AA;

Query Match 47.4%; Score 55; DB 32; Length 311;
Best Local Similarity 28.6%; Pred. No. 1.55e+02;
Matches 4; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

DB 190 lqrmepkmtwra 203
||| : :|||
QY 1 lkamdpplwkt 14

RESULT 10
ID W20095 standard; Protein; 446 AA.
AC W20095;
DT 08-JUL-1997 (first entry)
DE H. pylori cytoplasmic protein 10677187.aa.
KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
KM binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
OS Helicobacter pylori.
PN W09640893-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; U09122.
PR 07-JUN-1995; US-487032.
PR 01-APR-1996; US-630405.
PA (ASTR) ASTRA AB.
PI Berglund OT, Smith D, Mellgaard BL;
DR WPI; 97-052306/05.
DR N-PSDB; T67345.
PT Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PT infection, and to detect Helicobacter
PS Claim 61; Pages 329-330, 1481pp; English.
CC This sequence is a H. pylori cytoplasmic protein involved in
CC genome replication, transcription, recombination and repair.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds.
CC Useful as potential H. pylori life cycle activators or inhibitors.
CC The genomic sequence of H. pylori (ATCC 35679) was determined from
CC overlapping contigs generated by mechanically shearing the bacterial
CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
CC and the predicted coding regions defined by computer evaluation. To
CC identify likely H. pylori antigens for vaccine development, the amino
CC acid sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.
SQ Sequence 446 AA;

Query Match 47.4%; Score 55; DB 21; Length 446;
Best Local Similarity 50.0%; Pred. No. 1.55e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DB 403 akeppnlw 412
||| : :|||
QY 3 amdpplwkt 12

RESULT 11
ID R71376 standard; Protein; 551 AA.
AC R71376;
DT 21-NOV-1995 (first entry)
DE Human cystathionine beta-synthase.
KW Cystathionine; beta synthase; human; homocystinuria.
OS Homo sapiens.
PN W09507714-A.
PD 23-MAR-1995.
PF 12-SEP-1994; U10203.
PR 13-SEP-1993; US-120960.
PA (COLS) UNIV COLORADO.
PI Kraus JP;
DR WPI; 95-131185/17.
DR N-PSDB; O87430.
PT Purified DNA encoding human cystathionine beta-synthase - useful
PT for producing human cystathionine beta-synthase, used for
PS treating homocystinuria
PS Claim 7; Fig 1a-c; 45pp; English.

CC The amino acid sequence shown in R71376 is the human cystathionine
 CC beta-synthase (CBS). Human CBS can be used to treat patients with
 CC homocystinuria (CBS). The cDNA sequence, 087430, from which it is
 CC derived is also useful for screening CBS deficient patients for
 CC mutations in the CBS gene.
 SO Sequence 551 AA;

Query Match 47.4%; Score 55; DB 13; Length 551;
 Best Local Similarity 33.3%; Pred. No. 1.55e+02;
 Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 DB 36 keakeplvtrp 47
 OY 4 mdtptplwtkte 15

RESULT 12
 ID W21020 standard; protein; 594 AA.

AC W21020;
 DT 22-JUL-1997 (first entry)
 DE H. pylori cytoplasmic protein, hp0544orf18.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope; genome;
 KW replication; transcription; recombination; repair.
 OS Helicobacter pylori.
 PN W09640893-A1.
 PD 19-DEC-1996.
 PR 06-JUN-1996; U09122.
 PR 07-JUN-1995; US-487032.
 PR 01-APR-1996; US-630405.
 PA (ASTRA) ASTRA AB.
 PI Berzlinch OT, Smith D, Mellgaard BL;
 DR N-PSDB: T68273.
 PT Helicobacter pylori nucleic acid sequences and related
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
 PT infection, and to detect Helicobacter
 PS Claim 61; Pages 1410-1411; 1481pp; English.
 CC The present sequence is a Helicobacter pylori cytoplasmic protein
 CC involved in genomic replication, transcription, recombination and repair.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds, useful
 CC as potential H. pylori life cycle activators or inhibitors. The genomic
 CC sequence of H. pylori (ATCC 55679) was determined from overlapping
 CC contigs generated by mechanically shearing the bacterial DNA. The
 CC sequences were analysed for ORF of at least 180 nucleotides, and the
 CC predicted coding regions defined by computer evaluation. To identify
 CC likely H. pylori antigens for vaccine development, the amino acid
 CC sequences predicted from various ORF were analysed for significant
 CC homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 CC production, e.g. in E. coli hosts.
 SO Sequence 594 AA;

Query Match 47.4%; Score 55; DB 22; Length 594;
 Best Local Similarity 50.0%; Pred. No. 1.55e+02;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DB 451 akeppnlwv 460
 OY 3 amdtptplwtkte 12

RESULT 13
 ID R04197 standard; protein; 769 AA.
 AC R04197;
 DT 12-JUN-1990 (first entry)
 DE Env gene of simian immunodeficiency virus.
 KW SIV; simian immunodeficiency virus; AIDS; HIV; vaccine; ss.
 OS SIV.
 PN J01289485-A.

PD 21-NOV-1989.
 PR 16-JUN-1988; J19023.
 PR 16-JUN-1988; JP-119023.
 PA (ROFU) Tosa Nanyo Kogyo KK.
 DR WPI: 90-005176/01.
 DR N-PSDB: 002829.

PT DNA complementary to RNA of simian immuno-deficiency virus -
 PT used in vaccines and for diagnosis of AIDS.
 PS Disclosure: Fig 4; 6pp; Japanese.
 CC Env gene derived from RNA of SIV from which vector plasmid pSAH121
 CC may be constructed and incorporated into an E.coli vector. Useful in
 CC developing a vaccine against and in diagnosis of AIDS.
 SO Sequence 769 AA;

Query Match 47.4%; Score 55; DB 2; Length 769;
 Best Local Similarity 50.0%; Pred. No. 1.55e+02;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 DB 43 mtpstlwatn 54
 OY 4 mdtptplwtkte 15

RESULT 14
 ID W22986 standard; protein; 233 AA.

AC W22986;
 DT 08-OCT-1997 (first entry)
 DE Human serine protease 60 (SP60).
 KW Human; colon carcinoma; COLO 201; cell line; serine protease; SP60;
 KW screening; inhibitor; treatment; disease.
 OS Homo sapiens.
 PN J09149790-A.
 PD 10-JUN-1997.
 PR 24-JUL-1996; 212196.
 PR 29-SEP-1995; JP-275105.
 PA (SUNR) SONTORY LTD.
 DR WPI: 97-357902/33.
 DR N-PSDB: T79127.
 PT Human colon carcinoma derived serine protease(s) SP59, SP60 and SP67
 PT study agent for treatment of various diseases
 PT study agent for treatment of various diseases
 PS Claim 1; Pages 11-12; 16pp; Japanese.
 CC The present sequence is the human colon carcinoma COLO 201
 CC cell line derived serine protease 60 (SP60), which can be used to
 CC screen for specific inhibitors, e.g. to search for, or study an
 CC agent for the treatment of various diseases.
 SO Sequence 233 AA;

Query Match 46.6%; Score 54; DB 23; Length 233;
 Best Local Similarity 55.6%; Pred. No. 1.91e+02;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

DB 114 lpatcpplw 122
 OY 4 mdtptplwtkte 12

RESULT 15
 ID W56695 standard; protein; 396 AA.
 AC W56695;
 DT 24-JUL-1998 (first entry)
 DE Tetracycline resistance sequence contained in plasmid pTETL.
 KW Tns transposase; modified; enzyme; in vitro transposition; mutant;
 KW target; marker; transposon 5; plasmid pTETL; tetracycline resistance.
 OS Synthetic.
 OS Escherichia coli.
 PN W09810077-A1.
 PD 12-MAR-1998.
 PR 09-SEP-1997; U15941.
 PR 02-MAY-1997; US-850880.
 PR 09-SEP-1996; US-814877.
 PA (WISC) WISCONSIN ALUMNI RES FOUND.
 PI Goryshin IV, Reznikoff WS, Zhou H;

DR WPI: 98-193627/17.
 DR N-PSDB: V28398.
 PT Modified Tn5 transposase construct used in novel system for in vitro
 transposition - used to, e.g. create absolute defective mutants.
 PT provide selective markers and to facilitate insertion of specialised
 PT DNA sequences into target DNA
 PS Disclosure: Pages 35-46; 73pp; English.
 CC This tetracycline resistance sequence is contained in the plasmid
 CC pR2T1 which is used to demonstrate in vitro transposition of a
 CC transposable element located between a pair of Tn5 (transposon 5)
 CC outside end (OE) termini. The invention provides a genetic construct
 CC that contains a nucleotide sequence encoding a modified Tn5 transposase
 CC enzyme that has both greater avidity for Tn5 OE repeats and is less
 CC likely to assume an inactive multimeric form than a wild type Tn5
 CC transposase and a transposable DNA sequence flanked at its 5' and 3' ends
 CC by an 18 or 19 base pair flanking DNA sequence comprising nucleotide A
 CC at position 10, T at 11 and A at 12. The modified Tn5 transposase and
 CC the transposable DNA which is a DNA donor molecule are used in a system
 CC for in vitro transposition. The system and method can be used to create
 CC absolute defective mutants, to provide selective markers to target DNA,
 CC to provide portable regions of homology to a target DNA, to facilitate
 CC insertion of specialised DNA sequences into target DNA, to provide primer
 CC binding sites or tags for DNA sequencing, to facilitate production of
 CC genetic fusion for gene expression studies and protein domain mapping, as
 CC well as to bring together other desired combinations of DNA sequences
 CC (combinatorial genetics). The modified Tn5 transposase facilitates in
 CC vitro transposition reaction rates of at least about 100-fold higher
 CC than can be achieved using wild type transposase (as measure in vivo).
 CC In vitro transposition using this system can also use donor DNA and
 CC target DNA that is circular or linear. The system also requires no
 CC outside high energy source and no other protein other than the modified
 CC transposase.
 SQ Sequence 396 AA:

Query Match 46.6%; Score 54; DB 30; Length 396;
 Best Local Similarity 66.7%; Pred. No. 1.91e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 90 matpvlw1 98
 1 11 111
 QY 4 mdtpplw1 12

Search completed: Wed Dec 29 14:05:47 1999
 Job time : 29 secs.